Title
The signal of mantle anisotropy in the coupling of normal modes

Journal
Geophysical Journal International, 175(3)

ISSN
0956-540X

Authors
Beghein, Caroline
Resovsky, Joseph
Van Der Hilst, Robert D

Publication Date
2008-12-01

DOI
10.1111/j.1365-246x.2008.03970.x

Peer reviewed

Geophys. J. Int. (2008) **175,** 1209–1234 doi: 10.1111/j.1365-246X.2008.03970.x

The signal of mantle anisotropy in the coupling of normal modes

Caroline Beghein,[1] Joseph Resovsky[2] and Robert D. van der Hilst[3]

[1] *Earth and Space Sciences Department, University of California Los Angeles,* 595 *Charles Young Drive East,* Box 951567, *Los Angeles,* CA 90095-1567*, USA. E-mail: cbeghein@ucla.edu*
[2] *Roosevelt Academy, Science Department,* PO Box 94, 4330 *AB Middelburg, the Netherlands*
[3] *Massachusetts Institute of Technology,* 77 *Massachusetts avenue, building* 54-526, *Cambridge,* MA 02139*, USA*

Accepted 2008 September 10. Received 2008 September 9; in original form 2008 May 5

SUMMARY

We investigate whether the coupling of normal mode (NM) multiplets can help us constrain mantle anisotropy. We first derive explicit expressions of the generalized structure coefficients of coupled modes in terms of elastic coefficients, including the Love parameters describing radial anisotropy and the parameters describing azimuthal anisotropy (J_c, J_s, K_c, K_s, M_c, M_s, B_c, B_s, G_c, G_s, E_c, E_s, H_c, H_s, D_c and D_s). We detail the selection rules that describe which modes can couple together and which elastic parameters govern their coupling. We then focus on modes of type ${}_0S_l - {}_0T_{l+1}$ and determine whether they can be used to constrain mantle anisotropy. We show that they are sensitive to six elastic parameters describing azimuthal anisotropy, in addition to the two shear-wave elastic parameters L and N (i.e. V_{SV} and V_{SH}). We find that neither isotropic nor radially anisotropic mantle models can fully explain the observed degree two signal. We show that the NM signal that remains after correction for the effect of the crust and mantle radial anisotropy can be explained by the presence of azimuthal anisotropy in the upper mantle. Although the data favour locating azimuthal anisotropy below 400 km, its depth extent and distribution is still not well constrained by the data. Consideration of NM coupling can thus help constrain azimuthal anisotropy in the mantle, but joint analyses with surface-wave phase velocities is needed to reduce the parameter trade-offs and improve our constraints on the individual elastic parameters and the depth location of the azimuthal anisotropy.

Key words: Probability distributions; Surface waves and free oscillations; Seismic anisotropy.

1 INTRODUCTION

Seismic anisotropy, that is, the dependence of seismic wave velocities on the direction of propagation or polarization of the waves, has been found in several regions and at different depths inside the Earth. In many places, it is believed to be due to the preferred alignment of intrinsically anisotropic minerals (Karato 1988), but it could also be due to the layering of isotropic materials with contrasting elastic properties (Kendall & Silver 1996). Large-scale deformation processes have to be involved in the alignment or layering of minerals to observe seismic anisotropy at scales as large as several hundreds to thousands of kilometres. Accurate localization and characterization of seismic anisotropy can therefore improve our understanding of dynamic processes inside our planet.

Radial anisotropy (which is a case of transverse isotropy with the symmetry axis pointing in the radial direction) is required in the uppermost mantle to reconcile Love and Rayleigh wave dispersion (Anderson 1961), but its depth extent and how it varies in different tectonic settings is still somewhat unclear (Montagner & Tanimoto 1991; Gung *et al.* 2003; Beghein & Trampert 2004a; Panning & Romanowicz 2006). Radial anisotropy has also been reported for the transition zone, but the results are highly variable among the different studies (Beghein & Trampert 2004b; Panning & Romanowicz 2006). Azimuthal anisotropy occurs when seismic wave velocity changes with the azimuth of propagation. It was first observed by Hess (1964) in the Pacific ocean through the azimuthal dependence of propagation of P_n waves. Since then, it has been found at various places in the uppermost mantle and in the D″ layer, using various data such as shear wave splitting (e.g. Mitchell & Helmberger 1973; Silver & Chan 1991; Vinnik *et al.* 1992) and surface waves (Forsyth 1975; Montagner & Tanimoto 1990). It could also be present inside or just below the transition zone, as shown by mineral physics experiments (Kavner & Duffy 2001) and seismology (Fouch & Fischer 1996; Wookey *et al.* 2002, for shear-wave splitting analyses; Trampert & van Heijst 2002, for inversions of overtone Love wave phase velocity maps) .

Mantle seismic anisotropy is commonly inferred from shear-wave splitting or surface-wave dispersion measurements. On the one hand, body waves have a much lower radial than lateral resolution, which makes it difficult to locate the depth of origin of seismic anisotropy from shear-wave splitting measurements alone. On the other hand, surface wave dispersion is more sensitive to the depth

distribution of anisotropy. The normal modes (NMs) of Earth's free oscillations are sensitive to mantle structure as well, and their frequencies can be used to constrain anisotropy at greater depths than surface waves and with higher depth resolution than body waves. NM oscillations are degenerate in the case of a spherically symmetric non-rotating, elastic and isotropic Earth, but in the real Earth, the degeneracy is lifted and the modes 'split' due to the presence of 3-D heterogeneities, anisotropy, ellipticity and rotation. The frequencies of the modes can be measured and used to constrain the 3-D (anisotropic) structure of the Earth (see Dahlen & Tromp 1998, for details, and Section 2).

Most NM splitting measurements are performed in the self-coupling approximation, also called quasi-degenerate theory, where each multiplet is treated as isolated in the spectrum. Because of symmetry considerations, this approximation puts constraints only on even degrees of aspherical structure (Woodhouse & Dahlen 1978). Resovsky & Ritzwoller (1998) showed that isotropic mantle models can explain structure coefficients measured in the self-coupling approximation relatively well, and several mantle models have since been derived from such data (Resovsky & Ritzwoller 1999; Ishii & Tromp 1999; Masters *et al.* 2000; Beghein *et al.* 2002). Even though isolated mode structure coefficients can be explained by isotropic mantle models, Earth's free oscillations are compatible with the presence of radial anisotropy in the upper mantle and can provide constraints on large-scale radial anisotropy in the mantle (Panning & Romanowicz 2006).

Only a few researchers have investigated the effect of anisotropy on the coupling of NMs or surface waves. The signal of inner core anisotropy on coupling of core-sensitive NMs was recently analysed (Irving *et al.* 2008) using the wide-band coupling method of Deuss & Woodhouse (2001) and was found to have a potentially large effect. With a simple zonally symmetric and transversely isotropic uppermost mantle model, Park (1993) showed that as opposed to the coupling interaction between two long-period Rayleigh waves, long-period Love–Rayleigh coupling is stronger with azimuthal anisotropy than with radial anisotropy. In addition, using a similar synthetic model of anisotropy, Yu & Park (1993) demonstrated that smooth (that is at low spherical harmonic order) anisotropic structures are much more 'efficient' than isotropic structures in generating cross-type (Rayleigh–Love or spheroidal–toroidal) coupling of long-period surface waves. More recently, using synthetic tests and a spectral inversion, Oda (2005) showed that the coupling of Earth's NMs can be used to constrain the isotropic and anisotropic lateral structure of the Earth. However, no model of mantle anisotropy (radial or azimuthal) has ever been derived using 'real' (as opposed to synthetic) NM coupling data alone. The effect of inner core anisotropy, however, Resovsky & Ritzwoller (1995) incorporated intermultiplet coupling into splitting analyses by generalizing the spectral fitting technique previously employed for isolated mode multiplets (Ritzwoller *et al.* 1986). In addition to odd-degree structure coefficients, they measured a few spheroidal–toroidal (${}_nS_l - {}_nT_{l+1}$) multiplets, which are sensitive to even-degree structure (see Appendix B for details on the sensitivity of those modes). The odd-degree structure coefficients of their catalogue were used to derive an isotropic model of the whole mantle (Resovsky & Ritzwoller 1998), hereafter referred to as RR98, but the ${}_nS_l - {}_nT_{l+1}$ modes have not yet been used. These specific multiplets are sensitive to isotropic and anisotropic structure throughout most of the mantle (down to a depth of 2000 km), but most of the energy is situated above approximately 1200 km depth, and therefore, these data have the potential to constrain transition zone anisotropy.

In the first part of this paper, we show, through theoretical development, which type of structure can cause coupling between modes, and we give the selection rules that determine which modes couple. We then focus on the degree two splitting measurements of coupled modes ${}_nS_l - {}_nT_{l+1}$ made by Resovsky & Ritzwoller (1998). Besides rotation and ellipticity, such cross-type coupling between modes with angular orders l differing by 1 can occur due to the presence of (1) topography at internal discontinuities; (2) density anomalies; (3) isotropic or radially anisotropic shear-wave velocity anomalies and (4) azimuthally anisotropic structure (see Section 2 for details). We demonstrate that current isotropic models of the crust and the mantle (with topography at uppermost-mantle discontinuities) cannot explain the splitting measurements (which are corrected for the effect of rotation and ellipticity). To examine whether shear-wave radial anisotropy can explain the data, we test several models of upper-mantle radial anisotropy, inferred previously from either surface-wave overtone data (Beghein & Trampert 2004a,b) or from surface and body waveform data (Panning & Romanowicz 2006). This analysis shows that not all the NM generalized structure coefficients determined by Resovsky & Ritzwoller (1998) can be explained with radially anisotropic structure alone. Furthermore, we tried to fit part of the splitting measurements that remains after correcting for the effect of crustal structure and upper-mantle radial anisotropy with azimuthal anisotropy and to determine if any robust constraint on this type of anisotropy can be obtained from NM coupling data.

Using a singular value decomposition method (Matsu'ura & Hirata 1982), we then show that the problem of modelling azimuthal anisotropy with the available corrected NM data is highly ill-posed, which implies that the solution can strongly depend on initial model assumptions, even if the problem is linear (see Tarantola 1987; Matsu'ura & Hirata 1982, and Appendix A for more details). We used a model space search technique, based on forward modelling, to obtain insight into the class of acceptable solutions. The Neighbourhood Algorithm (NA; Sambridge 1999a,b) enables us to determine the entire family of models that satisfy the data and considerably reduces the risk of converging to a solution that is dictated by the initial model assumptions. This statistical approach to seismic tomography enables us to determine the most likely model and to test hypotheses (e.g. determine the likelihood of having azimuthal anisotropy below 400 km depth). In addition, the sampling of the whole model space (within selected boundaries), including the model null-space (i.e. the part of the model space not constrained by the data), yields more reliable estimate of model uncertainties and therefore of the robustness of the features observed than a least-squares inversion. Parameter trade-offs are directly available, making model resolution assessments more straightforward.

2 NORMAL MODE COUPLING EQUATIONS

The NMs of the Earth, or free oscillation multiplets, vibrate at frequencies that are specific to each mode and that depend on the internal structure of our planet. These modes are identified by the radial and angular orders (n, l) of their eigenfunctions, and each of the $2l + 1$ singlets composing the multiplet is characterized by an azimuthal index m. The free oscillations of a spherically symmetric, non-rotating, elastic and (transversely) isotropic earth have degenerate frequencies, that is, all singlets within one multiplet vibrate with the same eigenfrequency. Rotation, ellipticity and the presence of heterogeneities or anisotropy generate the splitting of

these multiplets. In this case, each singlet is characterized by a its own eigenfrequency. These singlet frequencies can be observed and measured on the seismic spectrum.

The splitting of two coupled mode multiplets $K = (n, l)$ and $K' = (n', l')$ is usually determined by a splitting matrix $H^{KK'}_{mm'}$ (Woodhouse 1980), which can be obtained by inversion of the spectrum. This splitting matrix is given by

$$H^{KK'}_{mm'} = D^{nn'l}_{m}\delta_{ll'}\delta_{mm'} + \sum_{s,t}\Gamma^{ll'}_{s}c^{t}_{s(KK')}, \tag{1}$$

where structural degree s varies between $|l - l'|$ and $l + l'$, and where $t = -s, \ldots, +s$. Furthermore, δ is the Kroneker symbol and D includes the effect of multiplet spacing, rotation and ellipticity. Specific selection rules determine which type of multiplets can couple through rotation or ellipticity (see Dahlen & Tromp 1998, for a review). Finally, Γ is a factor including geometric selection rules that determine which aspherical structure can cause modal coupling, and $c^{t}_{s(KK')}$ are the generalized structure coefficients that are linearly related to Earth's 3-D structure at spherical harmonic degree s and order t (Edmonds 1960). A general expression of the dependence of $c^{t}{}_{s(KK')}$ to perturbations of the fourth-order elastic tensor Λ is given by

$$c^{t}_{s(KK')} = \sum_{\alpha\beta\gamma\delta}\int_{0}^{a} K^{\alpha\beta\gamma\delta}_{s}(r)\delta\Lambda^{\alpha\beta\gamma\delta}_{st}(r)\,\mathrm{d}r, \tag{2}$$

where α, β, γ and δ are defined in the generalized spherical harmonics formulation of Phinney & Burridge (1973) and can be equal to 0, + 1 or −1. Because of the symmetry properties of the elastic tensor, only 21 terms corresponding $N = \alpha + \beta + \gamma + \delta = 0, \pm 1, \pm 2$ or ± 4 appear in eq. (2) (see Appendix C).

In the particular case of radial anisotropy, the contribution of the elastic parameters to the structure coefficients are given by the $N = 0$ terms of eq. (2) (Mochizuki 1986; Dahlen & Tromp 1998):

$$\begin{aligned} c^{t}_{s(KK')} = \int_{0}^{a} \Big[& \delta A^{t}_{s}(r)K^{(KK')}_{A}(r) + \delta C^{t}_{s}(r)K^{(KK')}_{C}(r) \\ & + \delta N^{t}_{s}(r)K^{(KK')}_{N}(r) + \delta L^{t}_{s}(r)K^{(KK')}_{L}(r) \\ & + \delta F^{t}_{s}(r)K^{(KK')}_{F}(r)\Big]\,\mathrm{d}r, \end{aligned} \tag{3}$$

with a is the radius of the Earth. Parameters A, C, N, L and F are the five independent elastic constants necessary to fully describe a radially anisotropic medium (Love 1927). Functions $K^{(KK')}_{\alpha}$ ($\alpha = A, C, N, L, F$) are partial derivatives, or sensitivity kernels, which characterize how each pair of modes averages Earth's structure. A general expression of these sensitivity kernels can be found in Mochizuki (1986) and Appendix B. Selection rules (see Appendices B and C for details) imply that degree s radially anisotropic (or isotropic) structure can couple two modes of the same type (two spheroidal modes ${}_{n}S_{l} - {}_{n'}S_{l'}$ or two toroidal modes ${}_{n}T_{l} - {}_{n'}T_{l'}$) only if $l + l' + s$ is even and can cause ${}_{n}S_{l} - {}_{n'}T_{l'}$ (cross-type) coupling if $l + l' + s$ is odd. When a toroidal mode is coupled to another mode (${}_{n}T_{l} - {}_{n'}T_{l'}$ or ${}_{n}S_{l} - {}_{n'}T_{l'}$), the coupling can be caused by the two shear-wave related elastic parameters L and N but not by P-wave anisotropy or P-wave velocity. In the case of spheroidal–spheroidal coupling, all five elastic parameters describing radial anisotropy are involved. Note that the same selection rules apply if we consider isotropic velocity anomalies since isotropy is a particular case of radial anisotropy with $A = C$, $L = N$ and $F = A - 2L$. Density anomalies $\delta\rho^{t}_{s}$ and topography at internal discontinuities can also influence mode coupling (Woodhouse 1980), in which case eq. (3) becomes

$$\begin{aligned} c^{t}_{s(KK')} = \int_{0}^{a} \Big[& \delta A^{t}_{s}(r)K^{(KK')}_{A}(r) + \delta C^{t}_{s}(r)K^{(KK')}_{C}(r) \\ & + \delta N^{t}_{s}(r)K^{(KK')}_{N}(r) + \delta L^{t}_{s}(r)K^{(KK')}_{L}(r) + \delta F^{t}_{s}(r)K^{(KK')}_{F}(r) \\ & + \delta\rho^{t}_{s}(r)K^{(KK')}_{\rho}(r)\Big]\,\mathrm{d}r + \sum_{d} h^{t}_{sd}B_{sd(KK')}r^{2}_{d}, \end{aligned} \tag{4}$$

where the so-called boundary factors B_d are known functions of radial eigenfunctions and h_d is the relative amplitude of the topography at the boundary d located at radius r_d.

In a more general case of anisotropy, the selection rules become more complicated, and 16 terms must be added to eq. (3). In Appendix C we demonstrate that the total splitting of the coupled modes can be divided into two main contributions: splitting due to lateral variations in radial anisotropy ($N = 0$) and splitting due to azimuthal anisotropy ($N \neq 0$). The contribution of radial anisotropy to structure coefficients is given by eq. (3). The contribution of azimuthal anisotropy takes different forms depending on the parity of $l + l' + s$ and on whether the coupling occurs between modes of the same type or between a spheroidal mode and a toroidal mode. At degree s, the contribution of azimuthal anisotropy follows these rules:

(1) ${}_{n}S_{l} - {}_{n'}S_{l'}$ and ${}_{n}T_{l} - {}_{n'}T_{l'}$ coupling occurs through parameters $(G_c)_{st}$, $(B_c)_{st}$, $(H_c)_{st}$, $(E_c)_{st}$, i $(E_s)_{st}$, i $(J_s)_{st}$, i $(M_s)_{st}$, i $(K_s)_{st}$ and i $(D_s)_{st}$ if $l + l' + s$ is even,

(2) ${}_{n}S_{l} - {}_{n'}S_{l'}$ and ${}_{n}T_{l} - {}_{n'}T_{l'}$ coupling occurs through parameters i $(G_s)_{st}$, i $(B_s)_{st}$, i $(H_s)_{st}$, $(J_c)_{st}$, $(M_c)_{st}$, $(K_c)_{st}$ and $(D_c)_{st}$ if $l + l' + s$ is odd,

(3) ${}_{n}S_{l} - {}_{n'}T_{l'}$ coupling occurs through parameters $(G_s)_{st}$, $(B_s)_{st}$, $(H_s)_{st}$, i $(J_c)_{st}$, i $(M_c)_{st}$, i $(K_c)_{st}$ and i $(D_c)_{st}$ if $l + l' + s$ is even,

(4) ${}_{n}S_{l} - {}_{n'}T_{l'}$ coupling occurs through parameters i $(G_c)_{st}$, i $(B_c)_{st}$, i $(H_c)_{st}$, $(E_s)_{st}$, i $(E_c)_{st}$, $(J_s)_{st}$, $(M_s)_{st}$, $(K_s)_{st}$ and $(D_s)_{st}$ if $l + l' + s$ is odd,

where the 16 elastic coefficients J_c, J_s, K_c, K_s, M_c, M_s, B_c, B_s, G_c, G_s, E_c, E_s, H_c, H_s, D_c and D_s are defined in Chen & Tromp (2007) in terms of the elastic tensor (see also appendix C for details). We see that, for a given type of coupling (like-type or cross-type), selection rules are governed by a set of elastic parameters that depends on the parity of $l + l' + s$. As shown by Park (1997), in a particular case of azimuthal anisotropy, these selection rules are a generalization of the familiar odd/even-parity selection rules for isotropic (or radially anisotropic) structures.

It is interesting to note the similarity between the dependence of NM coupling and surface- and body-wave phase speed propagation in terms of elastic parameters. For instance, the 2ψ azimuthal variation of surface- and body-wave phase speed is controlled by B_c, B_s, G_c, G_s, H_c and H_s, and their 4ψ variation is controlled by E_c and E_s. These elastic coefficients also appear in the $|N| = 2$ and $|N| = 4$ parts of eq. (2), respectively. In addition to the 2ψ and 4ψ terms, body-wave phase speed is also determined by a ψ term, dependent on J_c, J_s, K_c, K_s, M_c and M_s, and a 3ψ term, dependent on D_c and D_s. There is thus a correspondence between the ψ terms of body wave propagation and the $|N| = 1$ terms of mode coupling and between the 3ψ terms of body wave propagation and the $|N| = 3$ terms of mode coupling. Table 1 compares and summarizes which of the 21 elastic coefficients affects surface-wave phase speed, body wave speed and NM

Table 1. Comparison of the effect of the 21 elastic coefficients on the propagation of body waves, the phase speed of surface waves and the coupling of normal modes. A cross 'X' indicates coefficients that affect mode coupling or wave propagation speed. A '0' indicates no effect. The conditions required for normal mode coupling are also given. For body wave speed, the notation of Chen & Tromp (2007) is used, with B_{33} governing compressional-wave phase speed and B_{11}, B_{22} and B_{12} governing shear-wave phase speed.

	B_{33}	B_{11}	B_{22}	B_{12}	Rayleigh	Love	${}_nS_l - {}_{n'}S_{l'}$	${}_nT_l - {}_{n'}T_{l'}$	${}_nS_l - {}_{n'}T_{l'}$
δA	x	x	0	0	x	0	$l+l'+s$ even	0	0
δC	x	x	0	0	x	0	$l+l'+s$ even	0	0
δL	x	x	x	0	x	x	$l+l'+s$ even	$l+l'+s$ even	$l+l'+s$ odd
δN	0	0	x	0	x	x	$l+l'+s$ even	$l+l'+s$ even	$l+l'+s$ odd
δF	x	x	0	0	x	0	$l+l'+s$ even	0	0
J_c	x	0	0	0	0	0	$l+l'+s$ odd	0	$l+l'+s$ even
J_s	x	0	0	0	0	0	$l+l'+s$ even	0	$l+l'+s$ odd
K_c	x	x	0	x	0	0	$l+l'+s$ odd	0	$l+l'+s$ even
K_s	x	x	0	x	0	0	$l+l'+s$ even	0	$l+l'+s$ odd
M_c	0	0	x	x	0	0	$l+l'+s$ odd	$l+l'+s$ odd	$l+l'+s$ even
M_s	0	0	x	x	0	0	$l+l'+s$ even	$l+l'+s$ even	$l+l'+s$ odd
G_c	x	x	x	x	x	x	$l+l'+s$ even	$l+l'+s$ even	$l+l'+s$ odd
G_s	x	x	x	x	x	x	$l+l'+s$ odd	$l+l'+s$ odd	$l+l'+s$ even
B_c	x	x	0	x	x	0	$l+l'+s$ even	0	$l+l'+s$ odd
B_s	x	x	0	x	x	0	$l+l'+s$ odd	0	$l+l'+s$ even
H_c	x	x	0	x	x	0	$l+l'+s$ even	0	$l+l'+s$ odd
H_s	x	x	0	x	x	0	$l+l'+s$ odd	0	$l+l'+s$ even
D_c	x	x	x	x	0	0	$l+l'+s$ odd	$l+l'+s$ odd	$l+l'+s$ even
D_s	x	x	x	x	0	0	$l+l'+s$ even	$l+l'+s$ even	$l+l'+s$ odd
E_c	x	x	x	x	x	x	$l+l'+s$ even	0	$l+l'+s$ odd
E_s	x	x	x	x	x	x	$l+l'+s$ even	0	$l+l'+s$ odd

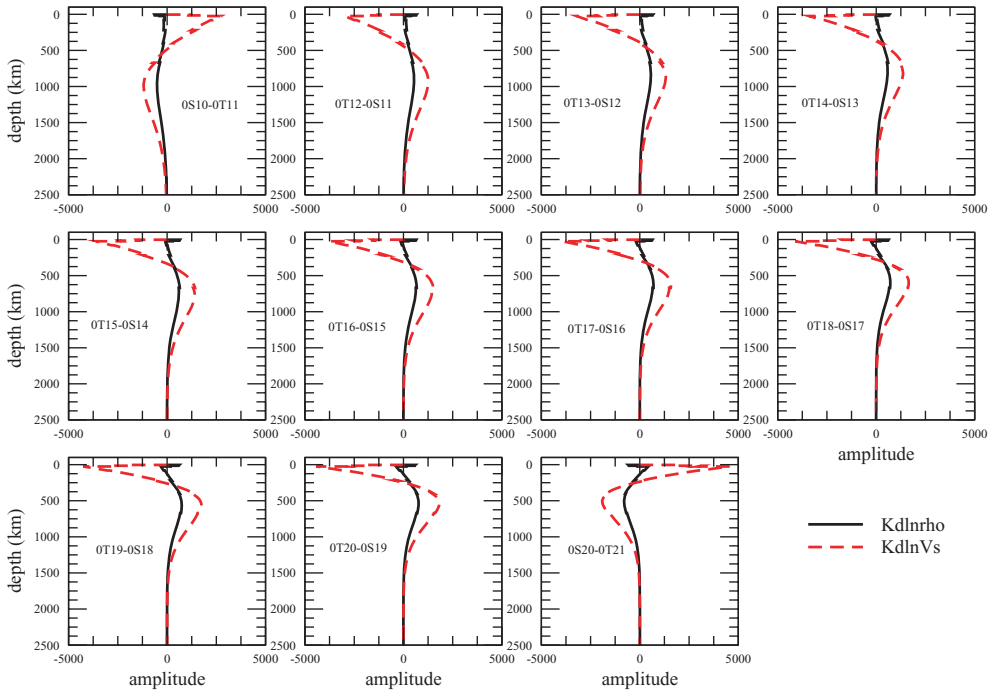

Figure 1. Sensitivity of coupled modes ${}_nS_l - {}_nT_{l+1}$ to relative shear wave and density anomalies. Those kernels are computed for structure coefficients expressed in μHz.

coupling structure coefficients, similarly to table 1 of Chen & Tromp (2007).

3 DATA

In this paper, we employ the degree two structure coefficients that were determined by Resovsky & Ritzwoller (1998) for modes ${}_0S_l - {}_0T_{l+1}$. These structure coefficients have never been included in modelling of Earth's interior for reasons explained below, but they have the potential of constraining structure at least as deep as the transition zone. As demonstrated in Section 2, they are sensitive to the two shear wave related elastic parameters (L and N), density and six of the 21 elastic parameters describing azimuthal anisotropy: G_c, B_c, H_c, K_s, M_s and J_s. The partial derivatives of these modes to density and isotropic shear-wave velocity anomalies are shown in Fig. 1. Fig. 2 displays the sensitivity to perturbations in elastic

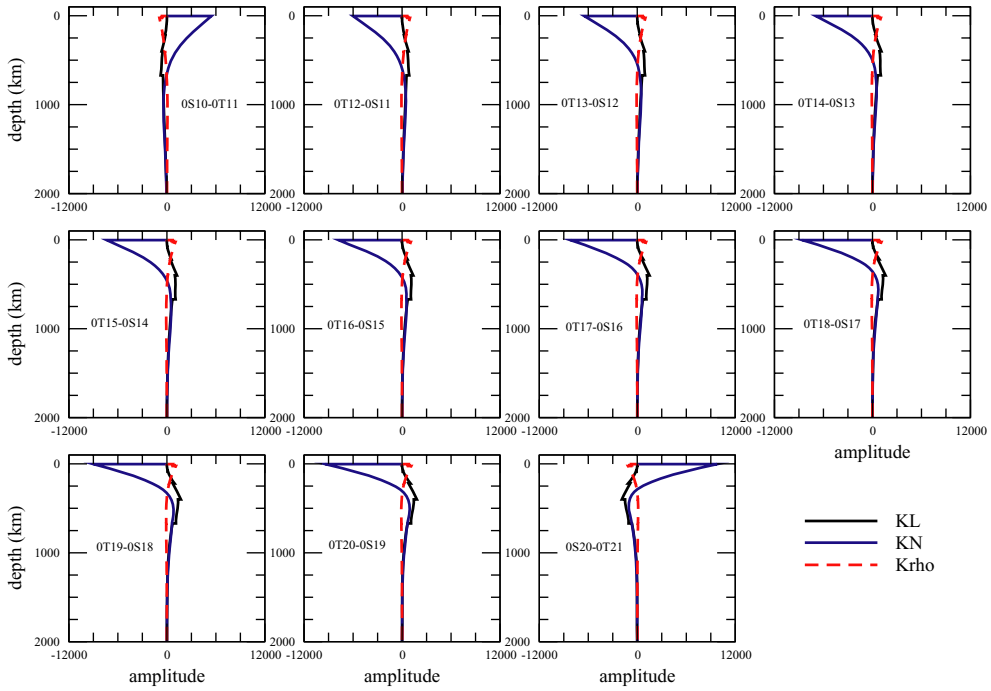

Figure 2. Sensitivity of coupled modes ${}_nS_l - {}_nT_{l+1}$ to elastic parameters L and N and to density anomalies. Those kernels are computed for structure coefficients expressed in μHz.

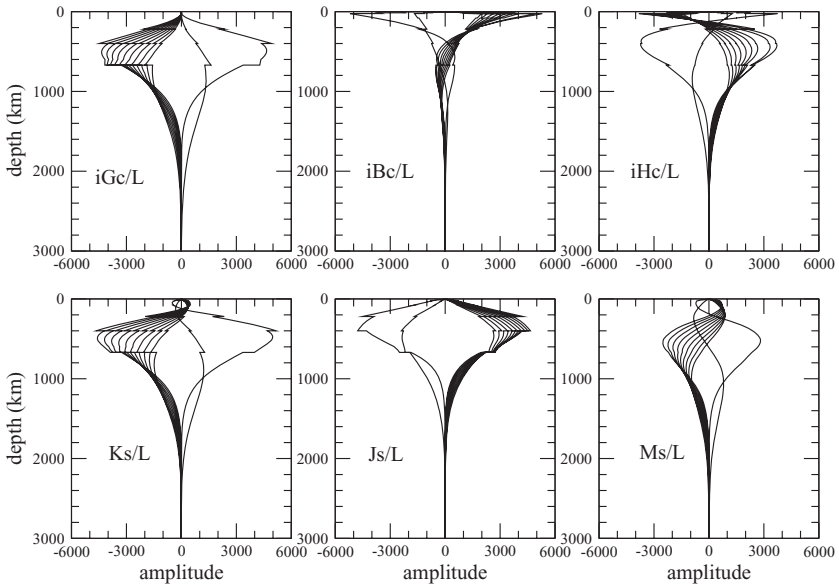

Figure 3. Sensitivity of coupled modes ${}_nS_l - {}_nT_{l+1}$ to six elastic parameters describing azimuthal anisotropy. The 11 different curves in each panel correspond to the 11 ${}_nS_l - {}_nT_{l+1}$ modes measured by Resovsky & Ritzwoller (1998) and analysed in this paper.

parameters L and N and in density, and Fig. 3 represents the sensitivity kernels for azimuthal anisotropy. We can see that the sensitivity of ${}_0S_l - {}_0T_{l+1}$ multiplets is most pronounced in the upper mantle, and that in the case of azimuthal anisotropy, the sensitivity is high in the crust for parameters B_c and H_c but almost inexistant for other parameters. Overal, NMs are more sensitive to isotropic or radially anisotropic structure, but their sensitivity to azimuthal anisotropy is not negligible.

The structure coefficients that are employed in this paper were determined by Resovsky & Ritzwoller (1998) as part of a detailed characterization of the free oscillation spectrum below 3 mHz. The spectrum was divided into windows around groups of closely spaced multiplets. Within such a group, the spacing between multiplet degenerate frequencies is less than 25 μHz, whereas spacing between groups is greater than 25 μHz, unless the groups could be separated using attenuation. Within each window, the splitting and coupling coefficients that describe the observed mode frequencies and attenuations were estimated using iterative least-squares regressions. The coefficients retrieved from such regressions were subjected to rigorous error control and assessment, designed to enhance their reliability as constraints on mantle structure. As possible data biases are of particular relevance to the analysis of a signal such as that of

mantle anisotropy, these procedures for error and bias control are worth reviewing.

First, all original seismic data were hand-edited and culled to optimize signal-to-noise ratios. Second, regressions were performed for the highest structural degrees that produced notable improvements in data fit, but the results for these highest degrees were consistently discarded because both the nature of the regression covariance matrices and synthetic experiments indicated that the signal of unmodelled coupling (from degrees too high to be estimated with the limited data set available at the time) is most likely to be absorbed by these high-degree coefficients. It seems reasonable to assume that this procedure also minimizes the impact of unmodelled broad-band coupling (Deuss & Woodhouse 2001).

Third, uncertainties were assigned to each structure coefficient using Monte-Carlo simulations, based on observations of the error processes expressed in the data themselves. Seismic noise in the frequency domain was measured in the gaps between multiplets for each seismogram and random perturbations were added to the synthetic seismograms to quantify the impact of these errors on the structure coefficient estimates.

Lastly, all coefficients and their associated uncertainty estimates were checked for a reasonable degree of consistency along overtone branches and with existing models. The differences between corresponding structure coefficients of multiplets with similar depth sensitivity, and the differences between the new coefficients and the predictions of some of the best-established mantle models at the time, were expected to be comparable to the differences between those models. Therefore, when differences that were not explained by the original uncertainty estimates were observed, those estimates were enlarged. In most cases, the uncertainties were already consistent (neither too small nor unreasonably large) with the observed differences, which increased confidence in the error modelling process described above. Measurements associated with large differences were omitted from the original catalogue (Resovsky & Ritzwoller 1998).

The spheroidal–toroidal structural coupling coefficients (${}_0S_l - {}_0T_{l+1}$) under consideration in this paper were almost excluded from the published catalogue. The degree 2 coefficients differed greatly from the predictions of existing isotropic mantle models, as described in Section 5.1, and (at the time) it was not believed that anisotropic structure in the upper mantle below the crust could be strong enough to explain such a signal. In the end, the coefficients were included in the catalogue because (1) they were retrieved using the same procedure that had yielded spheroidal–toroidal coupling coefficients with good model fit at odd degrees (e.g. ${}_5S_4 - {}_2T_4$); (2) the degree 4 and 6 coupling coefficients from the multiplet pairs were not unusual; (3) they displayed strong along-branch consistency; (4) their inclusion in regressions and synthetics provided significant improvements in data fit and (5) it was expected that crustal corrections would account for much of the unexplained signal from these multiplets. The latter did not prove to be the case.

4 APPROACH AND MOTIVATION

Modelling seismic anisotropy accurately can be a challenge. For instance, it is difficult to locate the depth of origin of seismic anisotropy from shear-wave splitting measurements because body waves lack vertical resolution. Surface waves and Earth's free oscillations have better depth resolution, but trade-offs exist between the different elastic parameters they are sensitive to. In addition, the uneven data coverage on Earth implies that seismic anisotropy cannot be constrained uniquely by seismology. This ill-posedness, or non-uniqueness of the solution, is inherent to most geophysical inverse problems. As pointed by Matsu'ura & Hirata (1982), in highly underdetermined cases the solution depends strongly on the initial assumptions made about the model parameters (parametrization, regularization, physical *a priori* constraints, distribution of data noise, *a priori* model uncertainties, etc.), and this does not happen only in non-linear problems.

In theory, one can transform an ill-posed problem into a well-posed problem by introducing sufficient *a priori* information and then solve the equations using a least-squares inversion (Jackson 1979). In seismology this can be done by combining different data sets (Masters *et al.* 2000; Marone & Romanowicz 2007a) or by imposing *a priori* relationships between elastic parameters (Montagner & Tanimoto 1991; Gung *et al.* 2003; Panning & Romanowicz 2006). The remaining parameters can be determined with a least-squares inversion, but the solution is still not unique and has to be regularized to choose one solution among all the models that can satisfy the data (see Appendix A). The choice of the regularization is, however, subjective and not based on true physical information. Moreover, a regularization compromises between improving the data fit and staying relatively close to an *a priori* reference model (Trampert 1998). When combined with the presence of strong trade-offs, the choice of the regularization can lead to a biased interpretation of the model if the true resolution of the model is not known (e.g. the problem of modelling inner core anisotropy from NM data Beghein & Trampert 2003). This can be avoided with model space search techniques based on forward modelling.

Forward modelling constitutes an alternative to a least-squares inversion. A common misconception in geophysics is that this sort of model space search approach is unnecessary in linear inverse problems, and that the robustness of the results of a least-squares inversion can be tested simply by slightly changing the regularization or the parametrization. This is, however, not always true (Sambridge *et al.* 2006). Model space searches based on forward modelling are designed to address problems with multiple misfit minima. Geophysicists often associate linear problems with a single global minimum in the cost function, but many linear inverse problems are ill-posed with non-Gaussian cost functions. As pointed out by Tarantola (1987), in such cases multiple minima are not only possible but also likely. An extended discussion of this aspect of the inverse problem is provided in Appendix A.

We thus adopted the NA due to Sambridge (1999a,b), which is a forward modelling technique that enables the exploration of the family of models that satisfy the data. Sampling the model space, including the model null-space (i.e. the part of the model space not constrained by the data) reduces considerably the risk of converging to a solution that is strongly dependent on the initial assumptions. A posterior probability density function (PPDF), also called 1-D marginal or likelihood, is associated with each model parameter and enables us to get a more reliable estimate of model uncertainties and therefore of the robustness of the observed features. In addition, parameter trade-offs are directly available, which makes model resolution assessments more straightforward.

The model space search is governed by the misfit of the models generated. We estimated the goodness of fit using a χ-misfit defined as follows.

$$\chi = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\frac{\left(d_k^{\mathrm{pred}} - d_k^{\mathrm{obs}}\right)^2}{{\sigma_k}^2}}, \tag{5}$$

where N is the total number of data, $d_k^{\rm pred}$ represents the kth data prediction and $d_k^{\rm obs}$ is the kth observed data with associated uncertainty σ_k. Values of χ of 1 or less correspond to model predictions that fall within one standard deviation of the data.

5 RESULTS

5.1 Isotropic models

In Section 2, eq. (4), we showed that structure coefficients describing coupling of the type ${}_nS_l - {}_nT_{l+1}$ can be due to 3-D density and isotropic shear-wave velocity anomalies and to topography at internal discontinuities. We decided to test the data first against predictions calculated from several isotropic d V_S mantle models: (1) MM2L12D8 (Resovsky & Ritzwoller 1998), which was obtained from isolated NM measurements made with the same method as the data analysed here; (2) S20RTS (Ritsema *et al.* 1999), which was derived from Rayleigh wave dispersion data, body-wave traveltimes and normal-mode splitting data; (3) S16B30 (Masters *et al.* 1996) that was obtained with surface-wave phase velocity maps, free oscillation structure coefficients and long-period body-wave absolute and differential traveltimes; (4) S12WM13 (Su *et al.* 1994), which used traveltimes of *ScS–S*, *SS–S* and other absolute traveltime data and waveform data; and (5) SAW24B16 (Mégnin & Romanowicz 2000), a V_{SH} model derived from body, surface and higher modes waveforms. Density anomalies were assumed to be proportional to shear-wave velocity anomalies using d $ln\ \rho = 0.4$ d $ln\, V_S$. This assumption, which implies that thermal anomalies are dominant in the mantle, is likely invalid in the deep mantle (Su & Dziewonski 1997; Kennett *et al.* 1998; Ishii & Tromp 1999; Resovsky & Trampert 2003), but our data are not sensitive to density anomalies at these depths. To predictions obtained with these bulk mantle models, we added predictions calculated from the degree two component of isotropic crustal model Crust5.1 (Mooney *et al.* 1998), including bathymetry and topography at discontinuities inside the crust and at the Moho. In the case of SAW24B16, we replaced the Crust5.1 Moho and seafloor topography with values given in SAW24B16.

We can see in Fig. 4 that most of the structure coefficients for the ${}_nS_l - {}_nT_{l+1}$ modes cannot be explained by bulk isotropic velocity anomalies alone. Predictions differ very little from one model to the other, but in most cases, the predicted values fall far outside the data uncertainties, especially for c_{20}. The highest c_{20} misfit calculated is $\chi = 10$ and is obtained with model MM2L12D8. The lowest c_{20} misfit is obtained with the *SH*-based model SAW24B16 but is still very high with a value of 8.4. Table 2 gives the misfit values obtained using these isotropic models for the five degree 2 spherical harmonic coefficients. Only predictions for Im (c_{22}) are consistent with the observed values. The χ-misfit for this spherical harmonic coefficient is close to 1, which shows that almost all the data are explained within uncertainties or close to it.

5.2 Radial anisotropy

As explained in Section 2, an alternative to isotropic V_S anomalies to generate coupling between modes is shear-wave radial anisotropy. We can see in Fig. 2 that the sensitivity of these modes to perturbations in the two shear wave related elastic coefficients L and N is mostly concentrated above the bottom of the transition zone. Radial anisotropy is therefore a possible candidate to explain the data.

We tested this possibility with models of radial anisotropy obtained by Beghein & Trampert (2004a,b), hereafter referred to as BT04, and Panning & Romanowicz (2006)— SAW642an. The way these two models were obtained differs in several ways, including the parametrization, the data used and the method employed to solve the problem. SAW642an was obtained using three component surface and body waveform data and an inversion for structure and source parameters based on non-linear asymptotic coupling theory (Li & Romanowicz 1995). In addition, *a priori* proportionality factors were imposed between the best-resolved (*S*-wave related) parameters and the least well resolved parameters (density anomalies, *P*-wave velocity and anisotropy anomalies and anisotropic parameter η). The data used in BT04 were fundamental mode and

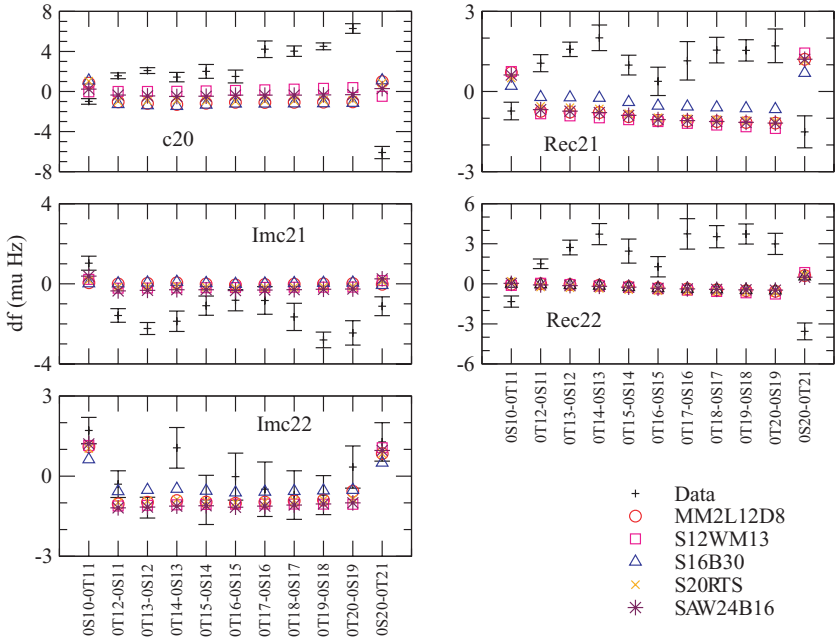

Figure 4. Observed splitting and splitting predicted by various isotropic mantle V_S and crustal models.

Table 2. χ-misfit obtained for various isotropic and anisotropic mantle models. The 'mean BT04' corresponds to predictions from the mean of the PPDFs associated with the BT04 models, that is, from the mean model. The most likely BT04 corresponds to the peak of the distributions of the BT04 models. The last column gives the misfit averaged over all five spherical harmonic coefficients.

	c_{20}	Re (c_{21})	Im (c_{21})	Re (c_{22})	Im (c_{22})	Average for all coefficients
MM2L12D8	10.0	5.4	4.1	4.5	1.17	5.0
S20RTS	9.3	5.1	3.8	4.7	1.16	4.8
S16B30	10.3	4.16	4.2	4.5	1.18	4.9
S12WM13	8.8	5.5	4.0	4.3	1.17	4.75
SAW24B16	8.4	5.3	3.5	4.65	1.3	4.6
Mean BT04	6.8	4.25	4.2	5.65	1.17	4.4
Most likely BT04	5.2	4.7	3.7	6.7	1.2	4.3
SAW642an	6.7	2.1	3.4	4.2	1.75	3.6
BT04 and azimuthal anisotropy	1.0	0.7	2.40	0.93	0.96	1.2

overtone surface-wave phase velocity maps at periods ranging from 50 to 150 s, with sensitivity down to the top of the lower mantle. No *a priori* scaling relationship was imposed between the model parameters, and the NA was employed to produce distributions of models for the five elastic parameters describing radial anisotropy and density anomalies instead of one single 'best-fitting' model.

Here, we used the degree two shear-wave velocity and anisotropy anomalies of SAW642an to predict the generalized structure coefficients $c^t_{s(KK')}$ of modes ${}_nS_l - {}_nT_{l+1}$, using the same prior relationship between d ρ and d V_S as used by Panning & Romanowicz (2006). We also calculated predictions for the $c^t_{s(KK')}$ using the family of BT04 models. We took advantage of the fact that a likelihood (or PPDF) is associated with the BT04 model parameters to obtain distributions of $c^t{}_{s(KK')}$ predictions, from which a mean prediction and standard deviation were extracted. In both cases, we included contributions from the crust using predictions calculated with Crust5.1. The predictions calculated with SAW642an and BT04 are plotted in Fig. 5 and misfit values are given in Table 2.

We find that the c_{20} structure coefficients are better explained with radial anisotropy than with isotropic models, and that the most likely BT04 model gives a better fit than SAW642an (Fig. 5 and Table 2). Nevertheless, the misfit is still high both with SAW642an and with BT04. Predictions for $Re(c_{21})$ improve quite significantly compared with predictions from isotropic models, especially with SAW642an, which fit the data better than any of the BT04 models. Im(c_{21}) and Re(c_{22}) remain largely unexplained with either model of radial anisotropy, and the fit stays close to the one obtained with isotropic structure. It is thus clear that we need to invoke an additional source of mode coupling to explain c_{20}, Re(c_{21}), Im(c_{21}) and Re(c_{22}) structure coefficients. Im(c_{22}) structure coefficients do not necessarily require any other source of mode coupling as their χ-misfit is close to 1 with radial anisotropy.

As explained above, SAW642an and BT04 were obtained using different methods. One of the differences is the use of prior constraints between the two *S*-wave related parameters and the least well-resolved parameters. To test what effect this has on the NM predictions, we applied the NA to the same degree two surface wave data used by Beghein & Trampert (2004a), and we imposed prior scaling relationships such as those used in SAW642an to find new distributions of shear-wave velocity and anisotropy models. New structure coefficients were then calculated (Fig. 6). The new range of predicted values generally overlaps with the BT04 predictions, but we found differences in the mean predicted $c^t_{s(KK')}$ compared with the mean predictions (Fig. 5a) and in their uncertainty. This uncertainty in the predicted data is larger with the BT04 models because model uncertainties are lower when *a priori* constraints are applied (Jackson & Matsu'ura 1985; Beghein 2008). Despite these small differences, the *a priori* proportionality factors cannot solely be responsible for discrepancies between predictions from the BT04 most likely model and from SAW642an since SAW642an predicted structure coefficients fall outside of the new model prediction range obtained with NA and *a priori* constraints.

5.3 Azimuthal anisotropy

5.3.1 Azimuthal anisotropy models

So far, we have neglected azimuthal anisotropy, but it is clear from currently available models that upper-mantle radial anisotropy alone cannot explain all the degree 2 ${}_nS_l - {}_nT_{l+1}$ mode data (Fig. 5). In this section, we examine whether azimuthal anisotropy can explain the remaining signal, after correction with predictions from an upper-mantle radial anisotropy model and a crustal isotropic model.

We seek to determine (1) a statistically robust ensemble of models of azimuthal anisotropy that could explain the remaining signal and (2) the likelihood that all this azimuthal anisotropy is located above the transition or that part of it lies deeper than 400 km depth. Because we have measurements for only 11 pairs of modes for each spherical harmonic coefficients and six elastic parameters to determine, this problem is ill-posed even if we employ a basic depth parametrization with only two layers. We opted for a two-layer parametrization with the following depth limits: 0–400 and 400–2000 km, 400 km delimiting the top of the transition zone. We cannot increase the number of layers any further and get results with meaningful misfit reduction because of the ill-posedness of the problem. A simple singular value decomposition, based on the method of Matsu'ura & Hirata (1982), showed that with these layers only 4 of the 12 unknown model parameters have significant singular values. This problem is thus clearly ill-conditionned and we decided to apply the NA to explore the ensemble of possible solutions and avoid converging to a solution strongly affected by the initial assumptions (Matsu'ura & Hirata 1982).

A first series of model space searches were performed on the mode data corrected with Crust5.1 and with the most likely BT04 model. We performed broad model space searches to reduce the risk of being trapped in a local minimum using well-chosen NA tuning parameters (see Sambridge 1999a, for details). For instance, for c_{20} at each iteration, 200 new models were generated within the 200 best-fitting Voronoi cells paving the model space; a total of 240 000 models were generated. We tested that the general features of these distributions are independent of the tuning parameters used, another sign that we are unlikely trapped in a local minimum. The 1-D marginal model distributions obtained for each spherical harmonic coefficient are shown in Fig. 7.

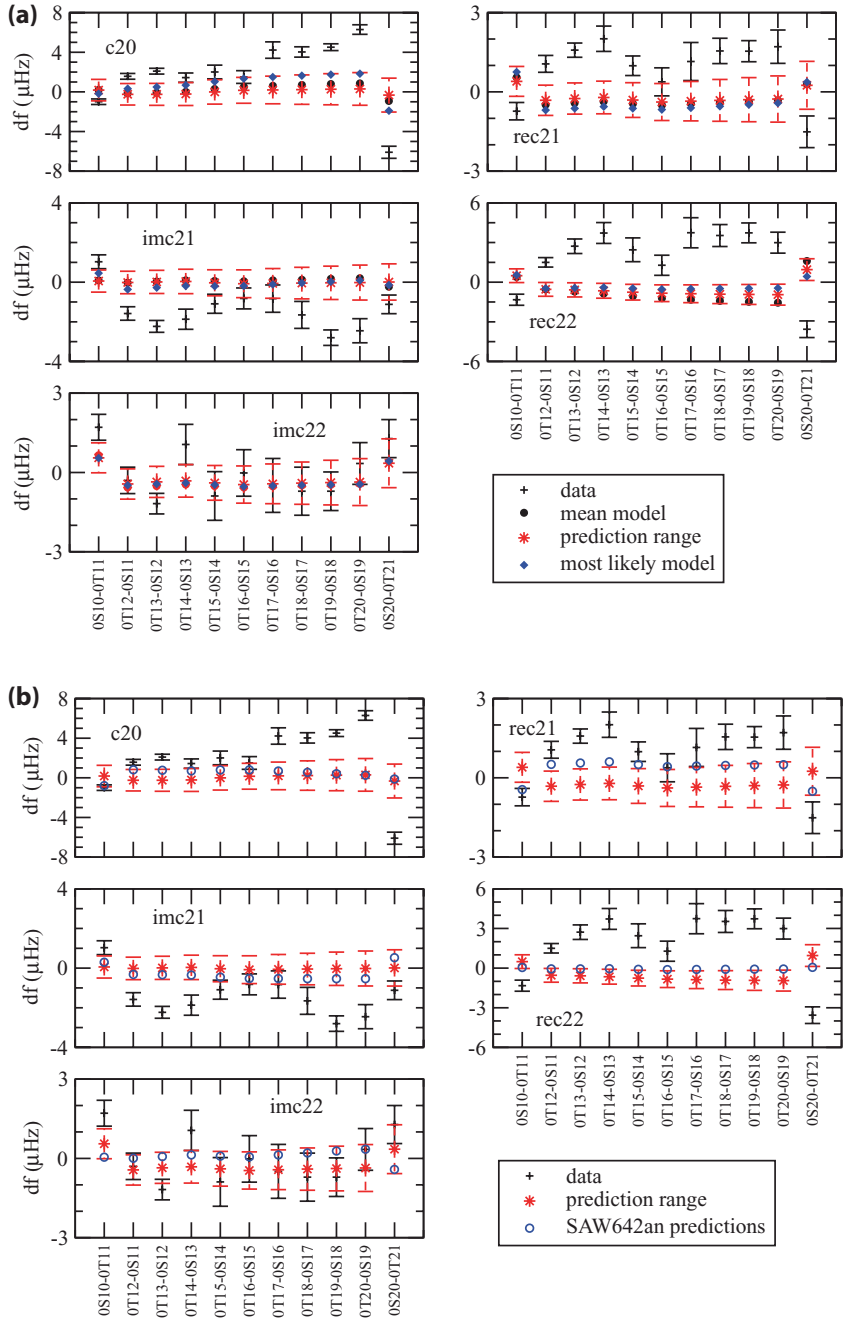

Figure 5. Observed data and range of data predictions computed using radially anisotropic models. Predictions from the distributions of BT04 radial anisotropy models are shown in (a). To obtain these predictions, we drew random values of dL_s^t, dN_s^t, and $d\rho_s^t$ from their PPDFs, and for each set of values we computed $c^t_{s(KK')}$ predictions. Thus, from the distributions of model parameters, we obtained a distribution of predictions, from which a mean and standard deviation were extracted. The mean and standard error are represented in red. Predictions from the mean radial anisotropic BT04 model and from the most likely radial anisotropic BT04 model are also displayed. Predictions from SAW642an and BT04 are shown in (b).

For c_{20}, the best fitting azimuthal anisotropy models explain the corrected data within uncertainties, with a χ-misfit of 1. In Fig. 7, we see that several elastic parameters are likely to depart from zero: in the top layer B_c, K_s and H_c have well-defined peaks, showing that B_c and K_s are most likely negative, and H_c is most likely positive. The PPDFs for the bottom layer imply that the corrected data favour the presence of azimuthal anisotropy deeper than 400 km: distributions for parameters G_c, J_s and M_s display clear peaks toward positive values, and K_s is most likely negative.

From Re(c_{21}), we find model distributions with most likely positive G_c and J_s and a likely negative K_s below depths of 400 km. In the top layer, however, the parameters do not show any clear signal: the width of their distributions is large, which could mean that either large trade-offs are present between the elastic parameters and/or that the elastic parameters vary vertically within our broad 400 km wide layer. The lowest χ misfit achieved for this spherical harmonic coefficient is 0.7.

For Im(c_{21}) peaks in the model distributions are clearly visible for a few elastic parameters, but none of the models generated can

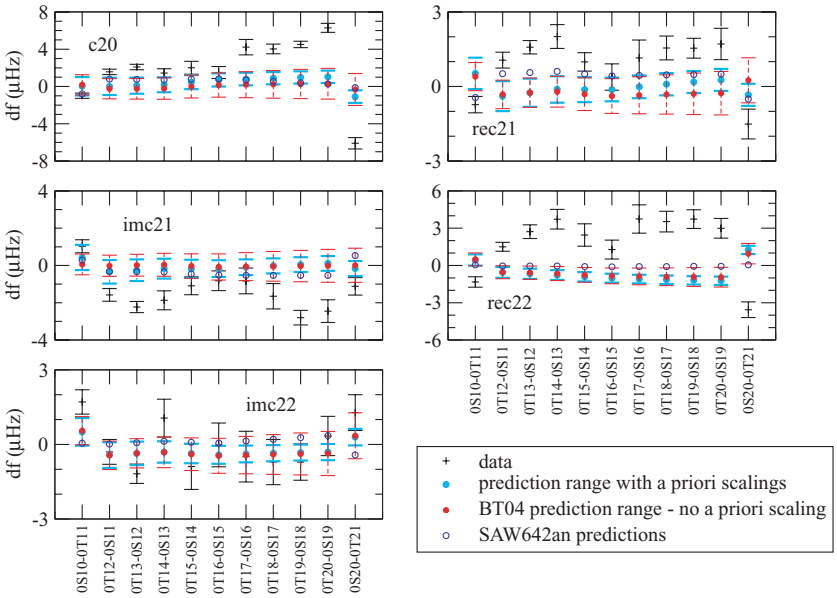

Figure 6. Observed data and range of data predictions computed using the distribution of models of radial anisotropy BT04 obtained by Beghein & Trampert (2004a,b) and using models obtained by forward modelling such as in BT04, but imposing *a priori* scaling relationships between anisotropic parameters as in SAW642an (Panning & Romanowicz 2006) . The distributions of models provides distributions of predictions from which a mean prediction and standard error were computed. Predictions from SAW642an are shown for comparison.

explain the corrected data better than within two standard deviations ($\chi = 2$). This is an improvement compared with the misfit obtained with radial anisotropy alone, but it is not sufficient to explain all the measurements. This could be due to contradictions among the measured mode structure coefficients or to more rapid depth variations of the anisotropy than accounted for with our depth parametrization, but we do not have enough data yet to test this.

NA applied to the Re(c_{22}) structure coefficients yielded best fitting models with $\chi = 0.93$. The PPDFs do not show clear peaks above 400 km, but H_c, J_s and K_s below 400 km display a likely departure from zero. We also applied NA to Im(c_{22}) even though radial anisotropy can already explain the measurements with $\chi = 1.2$. We found a slight improvement in the data fit by including azimuthal anisotropy, with a misfit going down to 0.96. The PPDFs are wide but display a peak toward positive values for M_s above 400 km and negative values for J_s and K_s below 400 km. Since the initial Im(c_{22}) misfit was already quite good, we do not rule out that this slight fit improvement is insignificant. If the measurement uncertainties were slightly underestimated by Resovsky & Ritzwoller (1998), we would be able fit the data within error with radial anisotropy alone.

5.3.2 *Effect of data corrections*

It is important to determine how the corrections applied to the data affect the azimuthal anisotropy model distributions. As demonstrated in Fig. 5, NM predictions from SAW642an can differ from the most likely BT04 model predictions. The choice of the model of radial anisotropy used to correct the data will therefore affect the residual data, which can affect the model of azimuthal anisotropy obtained with a traditional inverse technique. By using a statistical approach such as NA, however, we can determine whether there are any robust features in the azimuthal anisotropy, independent of the choice of the radial anisotropy correction. We thus applied the NA to the NM data corrected using Crust5.1 and SAW642an and compared the new distributions of azimuthal anisotropy models with those obtained with the Crust5.1 and BT04 corrections.

We find that the new models display similarities with, and can fit the data as well as, the models obtained using the BT04 corrections (Fig. 8). For instance, from c_{20} we obtain models for which B_c, H_c and K_s in the top layer and G_c, K_S and J_s in the bottom layer display clear peaks and the sign of these peaks is the same as in Fig. 7. On the contrary, some parameters that were not well constrained using the BT04 corrections now appear to display a better-defined peak (e.g. G_c in the top layer), and others that appeared well constrained with the BT04 corrections are now undetermined (e.g. M_s in the bottom layer). For Re(c_{21}), we see that G_c and J_s at depths larger than 400 km have the same PPDFs in both cases. Conclusions are less easy to draw for the elastic parameters obtained from Im(c_{21}). Like what we found using the BT04 corrections, however, no model can explain the Im(c_{21}) data better than within two standard deviations. For Re(c_{22}), K_s below 400 km is consistently more likely negative with either model of radial anisotropy. For Im(c_{22}), we can draw the same conclusions as when using the BT04 correction: we obtain similar model distributions but the fit improvement is very small and probably not significant since the initial misfit was already as low as 1.2.

5.3.3 *Depth dependence of the anisotropy*

Our results show that NM coupling data favour the presence of azimuthal anisotropy at depths greater than 400 km, independent of the upper-mantle radial anisotropy correction applied. A scenario where no azimuthal anisotropy is required by the data below that depth would be represented by PPDFs with peaks around zero for all six elastic parameters in the bottom layer or by PPDFs that are close to being flat if no clear signal was contained in the data (because of large data errors, large parameter trade-offs, contradictions in the data, etc). This is not the case here: we see (Figs 7 and 8) that the

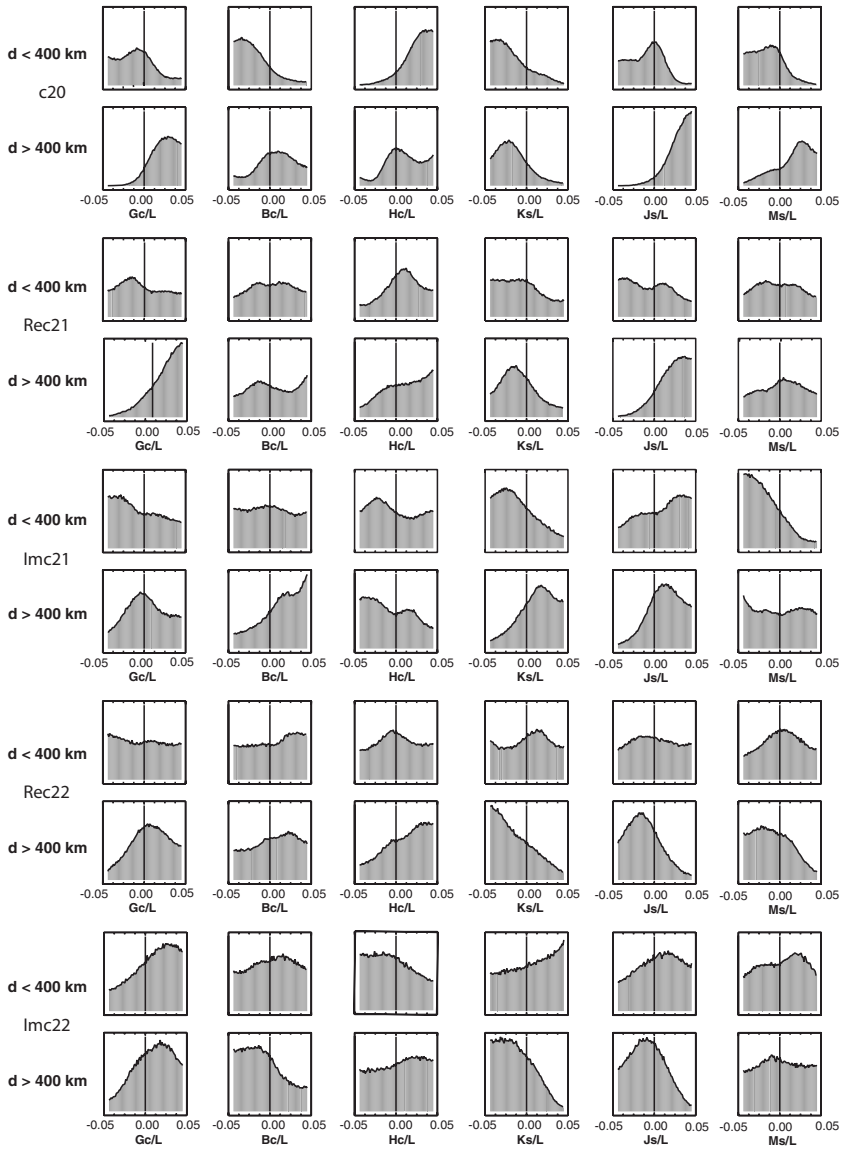

Figure 7. 1-D marginal distributions or likelihoods obtained by applying NA to the residual data, after correction for the effect of the crust and the most likely model of radial anisotropy BT04. We delimited the bounds of our model space to relative perturbations in elastic coefficients up to 5 per cent. We based our choice on the facts that (1) we should satisfy the conditions of application of perturbation theory, on which normal mode splitting equations are based, and that (2) if large-scale mantle azimuthal seismic anisotropy was much stronger, it would likely be more easily observable. The 1-D marginals represent the distributions of values for six elastic parameters describing azimuthal anisotropy in two layers: above 400 km depth and below 400 km depth. Parameters that have wide and close to flat distributions are not constrained by the data. Parameters for which the distribution has a clearly visible peak are much better constrained by the normal mode data. The vertical lines correspond to a model with no azimuthal anisotropy. The models obtained from the c_{20} measurements are shown in the top two rows, the results obtained from the Re (c_{21}) data are shown in the third and fourth rows, etc. For each *SH* coefficient, the top row displays the six elastic parameters in the top layer ($d < 400$ km), and the bottom row displays the six parameters averaged over our bottom layer ($d > 400$ km).

data are statistically better fitted with some amount of azimuthal anisotropy deeper than 400 km. Therefore, from the point of view of the data fit alone the likelihood of not having any azimuthal anisotropy below that depth is small. We need, however, to determine whether the data fit would change a lot if we did not allow for any azimuthal anisotropy below 400 km.

We thus performed another set of NA runs where we forced all the azimuthal anisotropy to reside in the upper 400 km of the mantle. To compare the new model misfits fairly with the misfits of the models shown in Fig. 7, we divided the upper 400 km in two layers, one from 0 to 200 and one from 200 to 400 km. We thus have 12 model parameters in both cases.

The c_{20} model distributions obtained for the upper 400 km are displayed in Fig. 9. These distributions are very wide, which means that the models are not well constrained due to a combination of data uncertainties, model parameter trade-offs, and possibly vertical variations within the layer chosen for our parametrization. Despite the large uncertainties, we see that G_c above 200 km has a distribution pointing toward negative values and the B_c likelihood in that same layer points toward a positive value. The lowest misfit obtained

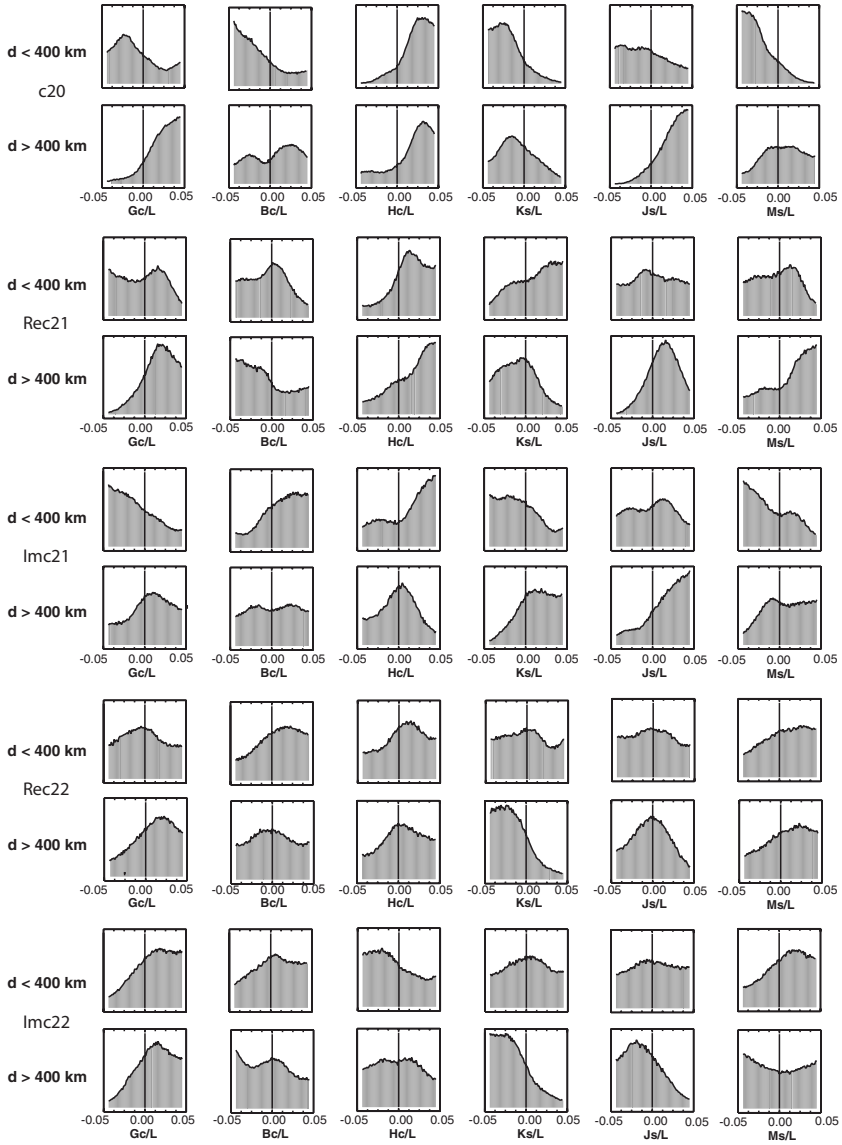

Figure 8. 1-D marginal distributions or likelihoods obtained by applying NA to the residual data, after correction for the effect of the crust and radial anisotropy using model SAW642an. See caption of Fig. 7 for details.

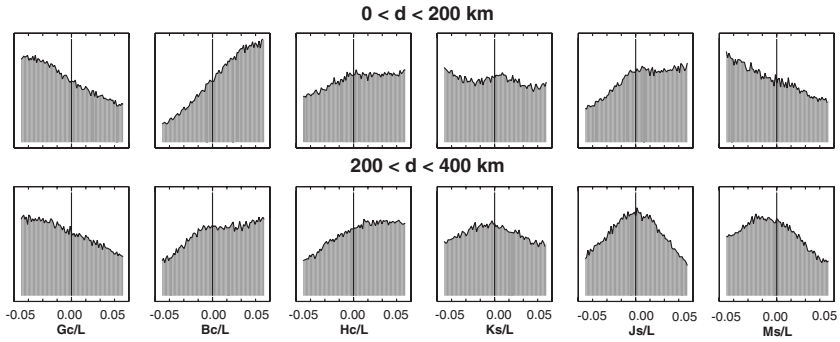

Figure 9. 1-D marginal distributions or likelihoods obtained by applying NA to the residual c_{20} data, after correction for the effect of the crust and radial anisotropy using most likely model BT04 and imposing the azimuthal anisotropy to be above a depth of 400 km.

was 1.3. It is a large improvement compared with the values obtained with radial anisotropy alone (χ around 7), and it is only slightly higher than what we obtained when we let the azimuthal anisotropy go below 400 km, in which case the lowest misfit reached was 1. From a statistical point of view, $\chi = 1.0$ is a better fit than $\chi = 1.3$, which explains why NA favours models with azimuthal anisotropy deeper than 400 km. From a practical point of view, we need to decide how much less likely are models that explain the

data within 1.3 times the standard deviation compared with models that explain the data within 1 standard deviation, but this is not a trivial decision to make. On the one hand, the data appear to contain a signal that lead to better-resolved model when we allow for deep anisotropy as the PPDFs present with better defined peaks than when we force the anisotropy to lie in the top 400 km of the mantle. On the other, we need to keep in mind that the normal data used and their error bars are coming from the inversion of seismic spectra, which itself involves some level of *a priori* information that might have influenced the structure coefficient estimates.

6 CONCLUSIONS

We showed that the degree two coupled mode structure coefficients, which were determined by Resovsky & Ritzwoller (1998) for $11\ {}_0S_l - {}_0T_{l+1}$ multiplets, cannot be explained by crustal structure and isotropic or radially anisotropic mantle models alone. Both model SAW642an and the radial anisotropy models obtained by Beghein & Trampert (2004a,b) were tested and none of them can explain the data. Analysis of the general equations relating coupled mode structure coefficients to anisotropy (Mochizuki 1986) shows that this type of coupled modes, at degree 2, is not only sensitive to shear-wave radial anisotropy but also to six elastic parameters describing azimuthal anisotropy. After correcting the data for the effect of structure in the crust and upper-mantle radial anisotropy, we showed that the remaining signal can be explained with azimuthal anisotropy.

Using a forward modelling technique, we established the probability that azimuthal anisotropy exists above and below depths of 400 km in the mantle. This conclusion is independent of the way the model space is sampled and of the radial anisotropy model used to correct the data. Not all the corrected NM structure coefficients can be explained by azimuthal anisotropy within error bars, but some show a large decrease in data misfit when azimuthal anisotropy is included.

The depth extent of the azimuthal anisotropy is, however, not yet well constrained. We tested the possibility of having all the azimuthal anisotropy confined within the top 400 km of the mantle. The models found after imposing such constraint have a slightly higher misfit than when anisotropy is allowed deeper than 400 km, and the likelihoods obtained do not display clear peaks. Therefore, from a purely statistical point of view, the data fit is better if the anisotropy goes below 400 km and the data seem to contain a robust signal at these depths. However, we cannot rule out yet, the possibility of azimuthal anisotropy confined above the transition zone. In addition, it should be noted that we cannot exclude that the crustal corrections applied in our analyses could affect the results as crustal corrections can have highly non-linear effects on NM and surface waves (Montagner & Jobert 1988; Boschi & Ekström 2002; Kustowski *et al.* 2007; Marone & Romanowicz 2007b; Bozdag & Trampert 2008). The way crustal corrections are made can be important, and 3-D variations in Moho depth affect not only the amount by which the measurements are corrected but also the partial derivatives used to retrieve structure at depth (e.g. Boschi & Ekström 2002; Marone & Romanowicz 2007b).

Our findings demonstrate that, like surface waves (Trampert & van Heijst 2002), coupled NM measurements require the presence of azimuthal anisotropy in the upper mantle. It is clear that NM coupling has the potential to constrain upper-mantle azimuthal anisotropy, but more data of that type are needed for future work if we want to better constrain the depth extent of the anisotropy. This could mean employing a full spectral coupling method such as the one developed by Deuss & Woodhouse (2001). In addition, since surface-wave phase velocities are sensitive to a subset of the elastic parameters controlling mode coupling, joint analyses of the two types of data will help break some of the parameter trade-offs and improve our constraints on the individual elastic parameters and the location at depth of the azimuthal anisotropy. This, in turn, could help us put new constraints on upper-mantle mineralogy and improve our understanding of mantle deformation.

ACKNOWLEDGMENTS

The authors thank Malcolm Sambridge for making his Neighbourhood Algorithm freely available. Comments by editors Mike Kendall and Gabi Laske, reviewer Barbara Romanowicz, Frank Kruger and two anonymous reviewers helped improve the manuscript.

REFERENCES

Anderson, D.L., 1961. Elastic wave propagation in layered anisotropic media, *J. geophys. Res.,* **66,** 2953–2963.

Beghein, C., 2008. Radial Anisotropy and a Priori Relationships Between Elastic Parameters: A Comparative Study, J. Geophys. Res., under revision.

Beghein, C. & Trampert, J., 2004a. Probability density functions for radial anisotropy from fundamental mode surface wave data and the neighbourhood algorithm, *Geophys. J. Int.,* **157**(3), 1163–1174.

Beghein, C. & Trampert, J., 2004b. Probability density functions for radial anisotropy: implications for the upper 1200 km of the mantle, *Earth planet. Sci. Lett.,* **217**(1–2), 151–162.

Beghein, C. & Trampert, J., 2003. Robust normal mode constraints on inner core anisotropy from model space search, science, **299,** 552–555. doi:10.1126/science.1078159.

Beghein, C., Resovsky, J. & Trampert, J., 2002. *P* and *S* tomography using normal mode and surface wave data with a neighbourhood algorithm, *Geophys. J. Int.,* **149**(3), 646–658.

Boschi, L. & Ekström, G., 2002. New images of the earths upper mantle from measurements of surface wave phase velocity anomalies, *J. geophys. Res.,* **107**(B4), 2059. doi:10.1029/2000JB000059.

Bozdag, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.,* **172,** 1066–1082.

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

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

Deuss, A. & Woodhouse, J.H., 2001. Theoretical free-oscillation spectra: the importance of wide band coupling, *Geophys. J. Int.,* **146,** 833–842.

Edmonds, A.R., 1960. *Angular Momentum and Quantum mechanics,* Princeton University Press, Princeton, NJ.

Forsyth, D.W., 1975. The early structural evolution and anisotropy of the oceanic upper mantle, *Geophys. J. R. astr. Soc.,* **43,** 103–162.

Fouch, M.J. & Fischer, K.M., 1996. Mantle anisotropy beneath northwest Pacific subduction zones, *J. geophys. Res.,* **101,** 15 987–16 002.

Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropic and the thickness of continents, *Nature,* **422,** 707–711.

Hess, H., 1964. Seismic anisotropy of the uppermost mantle under the oceans, *Nature,* **203,** 629–631.

Ishii, M. & Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of the Earth's mantle, *Science,* **285,** 1231–1236.

Irving, J.C.E., Deuss, A. & Andrews, J., 2008. Wide-band coupling of Earth's normal modes due to anisotropic inner core structure, *Geophys. J. Int.,* **174,** 919–929.

Jackson, D.-D., 1979. The use of a priori data to resolve non-uniqueness in linear inversion, *Geophys. J. R. astr. Soc.,* **57,** 137–147.

Jackson, D.-D. & Matsu'ura, M., 1985. A Bayesian approach to non-linear inversion, *J. geophys. Res.,* **90**(B1), 581–591.

Karato, S.-I., 1988. The role of recrystallisation in the preferred orientation of olivine, *Phys. Earth planet. Inter.,* **51,** 107–122.

Kavner, A. & Duffy, T.S., 2001. Strength and elasticity of ringwoodite at upper mantle pressures, *Geophys. Res. Lett.,* **28,** p. 2691.

Kendall, J. & Silver, P., 1996. Constraints from seismic anisotrophy on the nature of the lowermost mantle. *Nature,* **381,** 409–412.

Kennett, B.L.N., Widiyantoro, S. & van der Hilst, R.D., 1998. Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. geophys. Res.,* **103**(B6), 12 469–12 493.

Kustowski, B., Dziewonski, A.M. & Ekström, G., 2007. Nonlinear crustal corrections for normal-mode seismograms, *Bull. seism. Soc. Am.,* **97**(5), p. 1756–1762.

Li, X.D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121,** 695–709.

Love, A.E.H., 1927. *A Treatise on the Theory Mathematical of Elasticity,* Cambridge University Press, Cambridge, 643pp.

Marone, F. & Romanowicz, B., 2007a. The depth distribution of azimuthal anisotropy in the continental upper mantle, *Nature,* **447,** 198–203, doi:10.1038/nature05742.

Marone, F. & Romanowicz, B., 2007b. Non-linear crustal corrections in high-resolution regional waveform seismic tomography, *Geophys. J. Int.,* **170,** 460–467.

Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear velocity model of the mantle, *Phil. Trans. R. Soc. Lond., A,* **354,** 1385–1411.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A.M., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale, Seismology and Mineral Physics, Geophys. Monogr. Ser.,* Vol. 117, pp. 63–87, ed. Karato, S., AGU, Washington, DC.

Matsu'ura, M. & Hirata, N., 1982. Generalized least-squares solutions to quasi-linear inverse problems with *a priori* information, *J. Phys. Earth,* **30,** 451–468.

Mégnin, C. & Romanowicz, B., 2000. The shear velocity structure of the mantle from the inversion of body, surface and higher modes waveforms, *Geophys. J. Int.,* **143,** 709–728.

Mitchell, B.J. & Helmberger, D.V., 1973. Shear velocities at the base of the mantle from observations of S and ScS, *J. geophys. Res.,* **78**(26), 6009–6020.

Mochizuki, E., 1986. The free oscillations of an anisotropic and heterogeneous Earth, *Geophys. J. R. astr. Soc.,* **86,** 167–176.

Montagner, J.-P. & Jobert, N., 1988. Vectorial tomography, II: application to the indian ocean, *Geophys. J.,* **94,** 229–248.

Montagner, J.-P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.,* **95,** 4797–4819.

Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.,* **96**(B12), 20 337–20 351.

Montagner, J.-P., 1996. Surface waves on a global scale: influence of anisotropy and anelasticity, pp. 81–148, E. Boschi, G. Ekstrom, A. Morelli (Eds). Seismic Modelling of Earth Structure, Ed. Composition, Bologna, Italy.

Montagner, J.-P. & Guillot, L., 2000. Seismic anistropy in the Earth's Mantle. In Problems in Geophysics for the Next Millenium, PP. 218–253. eds Boschi, E., Ekström, G. Morelli, A. Editrice Compostri, Bologna, Italy.

Mooney, W., Laske, G. & Masters, G., 1998. Crust 5.1: a global crustal model at $5^\circ \times 5^\circ$, *J. geophys. Res.,* **103**(B1), 727–747.

Oda, H., 2005. An attempt to estimate isotropic and anisotropic lateral structure of the Earth by spectral inversion incorporating mixed coupling, *Geophys. J. Int.,* **160**(2), 667–682.

Panning, M. & Romanowicz, B., 2006. A three dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.,* **167,** 361–379.

Park, J., 1993. The sensitivity of seismic free oscillations to upper mantle anisotropy, 1: zonal symmetry, *J. geophys. Res.,* **98**(B11), 19 933–19 949.

Park, J., 1997. Free oscillations in an anisotropic earth: path-integral asymptotics, *Geophys. J. Int.,* **129,** 399–411.

Phinney, R.A. & Burridge, R., 1973. Representation of the elastic-gravitational excitation of a spherical Earth model by generalized spherical harmonics, *Geophys. J. R. astr. Soc.,* **34,** 451–487.

Resovsky, J.S. & Ritzwoller, M., 1995. Constraining odd-degree Earth structure with coupled free-oscillations, *Geophys. Res. Lett.,* **22**(16), 2301–2304.

Resovsky, J.S. & 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.S. & 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.S. & Trampert, J., 2003. Using probabilistic seismic tomography to test mantle velocity-density relationships, *Earth planet. Sci. Lett.,* **215,** 121–134.

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

Ritzwoller, M., Masters, G. & Gilbert, F., 1986. Observations of anomalous splitting and their interpretation in terms of aspherical structure, *J. geophys. Res.,* **91**(B10), 10 203–10 228.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm, I: searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm, II: appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Sambridge, M., Beghein, C., Simons, F.J. & Snieder, R., 2006. How do we understand and vizualize uncertainty?, *Leading Edge,* **25,** 542–546.

Silver, P.G. & Chan, W.W., 1991. Shear-wave splitting and subcontinental mantle deformation, *J. geophys. Res.,* **96,** 16 429–16 454.

Su, W.-J. & Dziewonski, A.M., 1997. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth planet. Int.,* **100,** 135–156.

Su, W.-J., Woodward, R.L. & Dziewonski, A.M., 1994. Degree 12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.,* **99**(B4), 6945–6980.

Tarantola, A., 1987. *Inverse Problem Theory, Methods for Data Fitting and Model Parameter Estimation,* Elsevier, Amsterdam.

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

Trampert, J. & van Heijst, H.J., 2002. Global azimuthal anisotropy in the transition zone, *Geophys. J. Int.,* **154**(1), 154–165.

Vinnik, L.P., Makeyeva, L.I., Milev, A. & Usenko, A.Y., 1992. Global patterns of azimuthal anisotropy and deformation in the continental mantle, *Geophys. J. Int.,* **111**(3), 433–447.

Woodhouse, J.H., 1980. The coupling and attenuation of nearly resonant multiplets in the Earth's free oscillation spectrum, *Geophys. J. R. astr. Soc.,* **61,** 261–283.

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.

Wookey, J., Kendall, J.-M. & Barruol, G., 2002. Mid-mantle deformation inferred from seismic anisotropy, *Nature,* **415,** 777–780.

Yu, Y. & Park, J., 1993. Upper mantle anisotropy and coupled-mode long-period surface waves, *Geophys. J. Int.,* **114,** 473–489.

APPENDIX A: INVERSE THEORY AND FORWARD MODELLING

The main advantage of using a forward modelling technique lies in the fact that we can determine the entire family of models that satisfy the data, without having to choose one model with a regularization. Interpretation of one model can be misleading and obtaining the entire family of models that satisfy a data set is a more robust and meaningful way of analysing and interpreting model features.
An inversion minimizes a cost function (e.g. a χ^2 misfit and/or a model norm term) to simultaneously optimize the data fit and use some prior information on the model space. A general form of the cost function is Tarantola (1987):

$$C_\lambda = \Delta_{\mathcal{D}}(\mathbf{d}, \mathbf{G}(\mathbf{m})) + \lambda \Delta_{\mathcal{M}}(\mathbf{m}, \mathbf{m}_0), \tag{A1}$$

where $\Delta_{\mathcal{D}}$ and $\Delta_{\mathcal{M}}$ are measures of the distance between observation $\mathbf{d}$ and prediction $\mathbf{G}(\mathbf{m})$ in the data space and between the solution $\mathbf{m}$ and a reference starting model $\mathbf{m}_0$ in the model space, respectively. A compromise between these two properties is reached by choosing an arbitrary value for the trade-off parameter λ. The choice of λ, of the *a priori* model $\mathbf{m}_0$, and of the model space and data space norms is arbitrary and constitutes some sort of *a priori* information.

It is well known that the cost function of non-linear inverse problems can have multiple minima, hence the solution can strongly depend on the prior information imposed. In that case, using a direct search approach is preferable to an inversion since it offers a way to obtain robust information on Earth's properties without having to introduce explicit *a priori* information on the model parameters for a given parametrization (i.e. $\lambda = 0$ and no explicit equation for $\Delta_{\mathcal{M}}$ is used). Forward modelling offers also a way to obtain reliable estimate of model resolution (that is parameter uncertainties and trade-offs). A common misconception in geophysics is that only non-linear problems have multiple minima and that the cost function of a linear inverse problem has only one global minimum. If it were true, identical solutions could be obtained with either a forward modelling technique or a traditional inverse method when the problem is linear. This is, however, not necessarily the case. When the problem is ill-posed, the cost function can be non-Gaussian, and as pointed out by Tarantola (1987), in such cases multiple minima are likely, even if the problem is linear.

Let us consider a linear inverse problem defined by $\mathbf{d} = \mathbf{G}\mathbf{m}$, where $\mathbf{d}$ represents the data, $\mathbf{m}$ represents the set of model parameters to determine and $\mathbf{G}$ represents a physical linear relation, relating model parameters and observables. The information we have on the observations can be represented by a volumetric probability $\rho_{\mathcal{D}}$ and the prior information we have on the model parameters is $\rho_{\mathcal{M}}$. Using the formulation of Tarantola (1987), the posterior information in the model space is given by the product of the *a priori* probability density in the model space $\rho_{\mathcal{M}}(\mathbf{m})$ by the probability density $\rho_{\mathcal{D}}(\mathbf{d})$ describing the result of the measurements:

$$\sigma_{\mathcal{M}}(\mathbf{m}) = \rho_{\mathcal{M}}(\mathbf{m})\rho_{\mathcal{D}}(\mathbf{G}\mathbf{m}) = \rho_{\mathcal{M}}(\mathbf{m})\rho_{\mathcal{D}}(\mathbf{d}). \tag{A2}$$

In many cases, data uncertainties are described by a Gaussian distribution

$$\rho_{\mathcal{D}}(\mathbf{d}) = \text{const} \times \exp\left[-\frac{1}{2}(\mathbf{d} - \mathbf{d}_{\text{obs}})^\dagger \mathbf{C}_{\mathcal{D}}^{-1}(\mathbf{d} - \mathbf{d}_{\text{obs}})\right], \tag{A3}$$

where $\dagger$ stands for the transpose of a matrix, const is a constant, $\mathbf{d}_{\text{obs}}$ represents the observations, and $\mathbf{C}_{\mathcal{D}}$ is the covariance matrix representing measurement uncertainties, which is often reduced to a diagonal matrix. We can further assume that the model parameters follow a Gaussian distribution centred on an *a priori* model $\mathbf{m}_0$ and with a model covariance matrix $\mathbf{C}_{\mathcal{M}}$:

$$\rho_{\mathcal{M}}(\mathbf{m}) = \text{const} \times \exp\left[-\frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^\dagger \mathbf{C}_{\mathcal{M}}^{-1}(\mathbf{m} - \mathbf{m}_0)\right]. \tag{A4}$$

In that case, the posterior model probability density function, or likelihood function, is given by

$$\sigma_{\mathcal{M}}(\mathbf{m}) = k \times \exp[-S(\mathbf{m})], \tag{A5}$$

where k is a constant and the misfit function $S(\mathbf{m})$ is

$$S(\mathbf{m}) = -\frac{1}{2}[(\mathbf{G}\mathbf{m} - \mathbf{d}_{\text{obs}})^\dagger \mathbf{C}_{\mathcal{D}}^{-1}(\mathbf{G}\mathbf{m} - \mathbf{d}_{\text{obs}}) + (\mathbf{m} - \mathbf{m}_0)^\dagger \mathbf{C}_{\mathcal{M}}^{-1}(\mathbf{m} - \mathbf{m}_0)]. \tag{A6}$$

Tarantola (1987) shows that $S(\mathbf{m})$ as given in eq. (A6) is Gaussian as well, with a mean given by

$$\tilde{\mathbf{m}} = \mathbf{m}_0 + \mathbf{C}_{\mathcal{M}}\mathbf{G}^\dagger \left(\mathbf{G}\mathbf{C}_{\mathcal{M}}\mathbf{G}^\dagger + \mathbf{C}_{\mathcal{D}}\right)^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{G}\mathbf{m}_0), \tag{A7}$$

and a resulting covariance matrix given by

$$\tilde{\mathbf{C}}_{\mathcal{M}} = \left(\mathbf{G}^\dagger \mathbf{C}_{\mathcal{D}}^{-1}\mathbf{G} + \mathbf{C}_{\mathcal{M}}^{-1}\right)^{-1}, \tag{A8}$$

$S(\mathbf{m})$ is Gaussian because the relationship $\mathbf{d} = \mathbf{G}(\mathbf{m})$ between data and model parameters is linear and because both the model parameter and the data error distributions are assumed to be Gaussian. If one of those three conditions is not met, the posterior model distribution is not necessarily Gaussian and the cost function can have multiple minima. As stated by Tarantola (1987), the further the relation $\mathbf{d} = \mathbf{G}(\mathbf{m})$ is from being linear, the further the posterior probability density $\sigma_{\mathcal{M}}(\mathbf{m})$ is from being a Gaussian. This is also true if either the data error or the model parameter distribution cannot be described by a Gaussian. Although in some cases it is possible to verify whether the data distribution is Gaussian (e.g. via numerous experiments), the probability distribution for the model parameters is usually not known *a priori*. It can be quite complex, especially in a high-dimensional model space, and the parameters may have non-standard (that is non-Gaussian) probability densities. By performing a model space exploration, one can determine model parameter distributions and find the whole ensemble of solutions that can explain the data, independent of prior explicit regularization and without assuming a prior Gaussian model distribution. Adopting a forward modelling approach can thus be advantageous when solving linear problems, as well as non-linear problems

APPENDIX B: GENERAL EQUATIONS

The splitting of two coupled mode multiplets $K = (n, l)$ and $K' = (n', l')$ is determined by a splitting matrix $H^{KK'}_{mm'}$ (e.g. Woodhouse 1980; Mochizuki 1986):

$$H^{KK'}_{mm'} = \frac{1}{2\omega I}\int\int\int \delta\Lambda_{ijkl}(E^*_{ij})_K(E_{kl})_{K'}\,\mathrm{d}V, \tag{B1}$$

where a sum over i, j, k, l is implicit. $\delta\Lambda_{ijkl}$ is a perturbation in an elastic constant, $\boldsymbol{E}$ is the strain tensor and E_{ij} is the strain tensor ij component. ω is the eigenfrequency of a particular multiplet for the zeroth-order isotropic Earth and I is a normalization factor. Integration is done over the volume of the Earth. The use of generalized spherical harmonics (Phinney & Burridge 1973) to express $\delta\Lambda_{ijkl}$ and the components of the strain tensor allows us to rewrite the last equation

$$H^{KK'}_{mm'} = \frac{1}{2\omega I}\sum_{s,t}(-1)^m 4\pi\gamma_l\gamma_{l'}\begin{pmatrix} l & s & l' \\ -m & t & m'\end{pmatrix}\sum_{\alpha\beta\gamma\delta}\int_0^a K_s^{\alpha\beta\gamma\delta}(r)\delta\Lambda_{st}^{\alpha\beta\gamma\delta}(r)\,\mathrm{d}r, \tag{B2}$$

with $\gamma = \sqrt{2l+1/4\pi}$, and where α, β, γ and δ are defined in the generalized spherical harmonics formulation and can be equal to 0, +1 or −1. We can divide the sum over α, β, γ, δ in different subsums. In the next sets of equations, we will write $\delta\Lambda_{st}^{\alpha\beta\gamma\delta}$ instead of $\delta\Lambda_{st}^{\alpha\beta\gamma\delta}(r)$, and we will be using $\Omega_0^l = \sqrt{l(l+1)/2}$ and $\Omega_2^l = \sqrt{(l+2)(l+1)/2}$. We will also be using Woodhouse's notation (Woodhouse 1980)

$$B^{N\pm}_{lsl'} = \frac{1}{2}\left[1\pm(-1)^{l+l'+s}\right]\sqrt{\frac{(l'+N)!(l+N)!}{(l'-N)!(l-N)!}}(-1)^N\begin{pmatrix} l & s & l' \\ -N & 0 & N\end{pmatrix}. \tag{B3}$$

If $N = \alpha+\beta+\gamma+\delta = 0$, we have five terms for $K_s^{\alpha\beta\gamma\delta}(r)\delta\Lambda_{st}^{\alpha\beta\gamma\delta}(r)$:

$$K_s^{0000}(r)\delta\Lambda_{st}^{0000} = r^2\dot{U}\dot{U}'B^{0+}_{lsl'}\delta\Lambda_{st}^{0000}, \tag{B4}$$

$$K_s^{++--}(r)\delta\Lambda_{st}^{++--} = \frac{1}{2}\left[(VV'+WW')B^{2+}_{lsl'} + (VW'-WV')\mathrm{i}B^{2-}_{lsl'}\right]\delta\Lambda_{st}^{++--}, \tag{B5}$$

$$K_s^{+-+-}(r)\delta\Lambda_{st}^{+-+-} = r^2FF'B^{0+}_{lsl'}\delta\Lambda_{st}^{+-+-}, \tag{B6}$$

$$K_s^{+-00}(r)\delta\Lambda_{st}^{+-00} = -r^2(F\dot{U}'+F'\dot{U})B^{0+}_{lsl'}\delta\Lambda_{st}^{+-00}, \tag{B7}$$

$$K_s^{+0-0}(r)\delta\Lambda_{st}^{+0-0} = -r^2\left[(XX'+ZZ')B^{1+}_{lsl'} + (XZ'-X'Z)\mathrm{i}B^{1-}_{lsl'}\right]\delta\Lambda_{st}^{+0-0}. \tag{B8}$$

U, V and W are the scalar eigenfunctions of multiplet (l, m) and F, Z and X are combinations of those eigenfunctions (see Woodhouse 1980, eq. A15). W and Z are non-zero for toroidal modes and *F, V, U* and *X* are non-zero for spheroidal modes. The dot represents the derivative with respect to radius r.

The case where $N = \alpha+\beta+\gamma+\delta = \pm1$, we have six terms in total:

$$K_s^{+000}(r)\delta\Lambda_{st}^{+000} = -r^2\left[\Omega_0^l\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}(X-\mathrm{i}Z)\dot{U}' + \Omega_0^{l'}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}(X'-\mathrm{i}Z')\dot{U}\right]\delta\Lambda_{st}^{+000}, \tag{B9}$$

$$K_s^{-000}(r)\delta\Lambda_{st}^{-000} = -r^2\left[\Omega_0^l\begin{pmatrix} l & s & l' \\ 1 & -1 & 0\end{pmatrix}(X+\mathrm{i}Z)\dot{U}' + \Omega_0^{l'}\begin{pmatrix} l & s & l' \\ 0 & -1 & 1\end{pmatrix}(X'+\mathrm{i}Z')\dot{U}\right]\delta\Lambda_{st}^{-000}, \tag{B10}$$

$$K_s^{++-0}(r)\delta\Lambda_{st}^{++-0} = -r\Omega_0^l\Omega_0^{l'}\left[\Omega_2^l\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}(V-\mathrm{i}W)(X'+\mathrm{i}Z') + \Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}(V'-\mathrm{i}W')(X+\mathrm{i}Z)\right]\delta\Lambda_{st}^{++-0}, \tag{B11}$$

$$K_s^{--+0}(r)\delta\Lambda_{st}^{--+0} = -r\Omega_0^l\Omega_0^{l'}\left[\Omega_2^l\begin{pmatrix} l & s & l' \\ 2 & -1 & -1\end{pmatrix}(V+\mathrm{i}W)(X'-\mathrm{i}Z') + \Omega_2^{l'}\begin{pmatrix} l & s & l' \\ -1 & -1 & 2\end{pmatrix}(V'+\mathrm{i}W')(X-\mathrm{i}Z)\right]\delta\Lambda_{st}^{--+0}, \tag{B12}$$

$$K_s^{+-+0}(r)\delta\Lambda_{st}^{+-+0} = r^2\left[\Omega_0^{l'}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}F(X'-\mathrm{i}Z') + \Omega_0^l\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}F'(X-\mathrm{i}Z)\right]\delta\Lambda_{st}^{+-+0}, \tag{B13}$$

$$K_s^{-+-0}(r)\delta\Lambda_{st}^{-+-0} = r^2\left[\Omega_0^{l'}\begin{pmatrix} l & s & l' \\ 0 & -1 & 1\end{pmatrix}F(X'+\mathrm{i}Z') + \Omega_0^l\begin{pmatrix} l & s & l' \\ 1 & -1 & 0\end{pmatrix}F'(X+\mathrm{i}Z)\right]\delta\Lambda_{st}^{-+-0}. \tag{B14}$$

The case where $N = \pm 2$, we have six terms as well.

$$K_s^{++00}(r)\delta\Lambda_{st}^{++00} = r\left[\Omega_0^l\Omega_2^l\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}(V - \mathrm{i}W)\dot{U}' + \Omega_0^{l'}\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\dot{U}(V' - \mathrm{i}W')\right]\delta\Lambda_{st}^{++00}, \tag{B15}$$

$$K_s^{--00}(r)\delta\Lambda_{st}^{--00} = r\left[\Omega_0^l\Omega_2^l\begin{pmatrix} l & s & l' \\ 2 & -2 & 0\end{pmatrix}(V + \mathrm{i}W)\dot{U}' + \Omega_0^{l'}\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 0 & -2 & 2\end{pmatrix}\dot{U}(V' + \mathrm{i}W')\right]\delta\Lambda_{st}^{--00}, \tag{B16}$$

$$K_s^{+0+0}(r)\delta\Lambda_{st}^{+0+0} = r^2\Omega_0^l\Omega_0^{l'}\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}\left[XX' - ZZ' - \mathrm{i}(ZX' + XZ')\right]\delta\Lambda_{st}^{+0+0}, \tag{B17}$$

$$K_s^{-0-0}(r)\delta\Lambda_{st}^{-0-0} = r^2\Omega_0^l\Omega_0^{l'}\begin{pmatrix} l & s & l' \\ 1 & -2 & 1\end{pmatrix}\left[XX' - ZZ' + \mathrm{i}(ZX' + XZ')\right]\delta\Lambda_{st}^{-0-0}. \tag{B18}$$

$$K_s^{+++-}(r)\delta\Lambda_{st}^{+++-} = -r\left[\Omega_0^l\Omega_2^l\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}F'(V - \mathrm{i}W) + \Omega_0^{l'}\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}F(V' - \mathrm{i}W')\right]\delta\Lambda_{st}^{+++-}, \tag{B19}$$

$$K_s^{---+}(r)\delta\Lambda_{st}^{---+} = -r\left[\Omega_0^l\Omega_2^l\begin{pmatrix} l & s & l' \\ 2 & -2 & 0\end{pmatrix}F'(V + \mathrm{i}W) + \Omega_0^{l'}\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 0 & -2 & 2\end{pmatrix}F(V' + \mathrm{i}W')\right]\delta\Lambda_{st}^{---+}. \tag{B20}$$

There are two terms corresponding to $N = \pm 3$ and two terms for $N = \pm 4$.

$$K_s^{+++0}(r)\delta\Lambda_{st}^{+++0} = -r\Omega_0^l\Omega_0^{l'}\left[\Omega_2^l\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}[VX' - WZ' - \mathrm{i}(WX' + VZ')\right] + \Omega_2^{l'}\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix} \times [V'X - W'Z - \mathrm{i}(W'X + V'Z)]\delta\Lambda_{st}^{+++0}, \tag{B21}$$

$$K_s^{---0}(r)\delta\Lambda_{st}^{---0} = -r\Omega_0^l\Omega_0^{l'}\left[\Omega_2^l\begin{pmatrix} l & s & l' \\ 2 & -3 & 1\end{pmatrix}[VX' - WZ' + \mathrm{i}(WX' + VZ')\right] + \Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 1 & -3 & 2\end{pmatrix} \times [V'X - W'Z + \mathrm{i}(W'X + V'Z)]\delta\Lambda_{st}^{---0}, \tag{B22}$$

$$K_s^{++++}(r)\delta\Lambda_{st}^{++++} = \Omega_0^l\Omega_0^{l'}\Omega_2^l\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ -2 & 4 & -2\end{pmatrix} \times [VV' - WW' - \mathrm{i}(WV' + VW')]\delta\Lambda_{st}^{++++}, \tag{B23}$$

$$K_s^{----}(r)\delta\Lambda_{st}^{----} = \Omega_0^l\Omega_0^{l'}\Omega_2^l\Omega_2^{l'}\begin{pmatrix} l & s & l' \\ 2 & -4 & 2\end{pmatrix} \times [VV' - WW' + \mathrm{i}(WV' + VW')]\delta\Lambda_{st}^{----}. \tag{B24}$$

Eqs (B9)–(B24) can be found in Mochizuki (1986).

B1 Like-type coupling

Here we give the equations for the coupling between two modes of the same type (spheroidal–spheroidal and toroidal–toroidal), and we make use of the symmetry properties of the 3j-Wigner symbols and of eq. (B3).

If $N = 0$, we have

$$K_s^{0000}(r)\delta\Lambda_{st}^{0000} = r^2\dot{U}\dot{U}'B_{lsl'}^{0+}\delta\Lambda_{st}^{0000}, \tag{B25}$$

$$K_s^{++--}(r)\delta\Lambda_{st}^{++--} = \frac{1}{2}(VV' + WW')B_{lsl'}^{2+}\delta\Lambda_{st}^{++--}, \tag{B26}$$

$$K_s^{+-+-}(r)\delta\Lambda_{st}^{+-+-} = r^2FF'B_{lsl'}^{0+}\delta\Lambda_{st}^{+-+-}, \tag{B27}$$

$$K_s^{+-00}(r)\delta\Lambda_{st}^{+-00} = -r^2(F\dot{U}' + F'\dot{U})B_{lsl'}^{0+}\delta\Lambda_{st}^{+-00}, \tag{B28}$$

$$K_s^{+0-0}(r)\delta\Lambda_{st}^{+0-0} = -r^2(XX' + ZZ')B_{lsl'}^{1+}\delta\Lambda_{st}^{+0-0}. \tag{B29}$$

For $|N| = 1$, we have the following terms

$$K_s^{+000}(r)\delta\Lambda_{st}^{+000} + K_s^{-000}(r)\delta\Lambda_{st}^{-000} = -r^2\left(\delta\Lambda_{st}^{+000} + (-1)^{l+l'+s}\delta\Lambda_{st}^{-000}\right)\left[\Omega_0^l X\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix} + \Omega_0^{l'}X'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}\right], \tag{B30}$$

$$K_s^{++-0}(r)\delta\Lambda_{st}^{++-0} + K_s^{-+0}(r)\delta\Lambda_{st}^{-+0} = -r\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{++-0} + (-1)^{l+l'+s}\delta\Lambda_{st}^{-+0}\right)\left[\Omega_2^l(VX' + WZ')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(V'X + W'Z)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right], \tag{B31}$$

$$K_s^{+-+0}(r)\delta\Lambda_{st}^{+-+0} + K_s^{-+-0}(r)\delta\Lambda_{st}^{-+-0} = r^2\left(\delta\Lambda_{st}^{+-+0} + (-1)^{l+l'+s}\delta\Lambda_{st}^{-+-0}\right)\left[\Omega_0^{l'}FX'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix} + \Omega_0^l F'X\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]. \tag{B32}$$

For $|N| = 2$:

$$\delta\Lambda_{st}^{++00}K_s^{++00} + \delta\Lambda_{st}^{-00}K_s^{-00} = r\left(\delta\Lambda_{st}^{++00} + (-1)^{l+l'+s}\delta\Lambda_{st}^{-00}\right)\left[\Omega_0^l\Omega_2^l V\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} + \Omega_2^{l'}\Omega_0^{l'}V'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{B33}$$

$$\delta\Lambda_{st}^{0+0+}K_s^{0+0+} + \delta\Lambda_{st}^{0-0-}K_s^{0-0-} = r^2\Omega_0^l\Omega_0^{l'}(XX' - ZZ')\left(\delta\Lambda_{st}^{0+0+} + (-1)^{l+l'+s}\delta\Lambda_{st}^{0-0-}\right)\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}, \tag{B34}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-} + \delta\Lambda_{st}^{---+}K_s^{---+} = -r(\delta\Lambda_{st}^{+++-} + (-1)^{l+l'+s}\delta\Lambda_{st}^{---+})\left[\Omega_0^l\Omega_2^l VF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} + \Omega_2^{l'}\Omega_0^{l'}V'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{B35}$$

For $|N| = 3$:

$$\delta\Lambda_{st}^{+++0}K_s^{+++0} + \delta\Lambda_{st}^{---0}K_s^{---0} = -r\Omega_0^{l'}\Omega_0^l\left(\delta\Lambda_{st}^{+++0} + (-1)^{l+l'+s}\delta\Lambda_{st}^{---0}\right)\left[\Omega_2^l(VX' - WZ')\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(V'X - W'Z)\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right]. \tag{B36}$$

Finally, for $|N| = 4$,

$$\delta\Lambda_{st}^{++++}K_s^{++++} + \delta\Lambda_{st}^{----}K_s^{----} = \left(\delta\Lambda_{st}^{++++} + (-1)^{l+l'+s}\delta\Lambda_{st}^{----}\right)\Omega_2^{l'}\Omega_0^{l'}\Omega_2^l\Omega_0^l(VV' - WW')\begin{pmatrix} l & s & l' \\ -2 & 4 & -2\end{pmatrix}. \tag{B37}$$

B2 Cross-type coupling

Here we give the equations for the coupling between a spheroidal mode and a toroidal mode.

If $N = 0$, we have:

$$K_s^{0000}(r)\delta\Lambda_{st}^{0000} = 0, \tag{B38}$$

$$K_s^{++--}(r)\delta\Lambda_{st}^{++--} = \frac{1}{2}\mathrm{i}(VW' - WV')B_{lsl'}^{2-}\delta\Lambda_{st}^{++--}, \tag{B39}$$

$$K_s^{+-+-}(r)\delta\Lambda_{st}^{+-+-} = 0, \tag{B40}$$

$$K_s^{+-00}(r)\delta\Lambda_{st}^{+-00} = 0, \tag{B41}$$

$$K_s^{+0-0}(r)\delta\Lambda_{st}^{+0-0} = -r^2\mathrm{i}(XZ' - X'Z)B_{lsl'}^{1-}\delta\Lambda_{st}^{+0-0}, \tag{B42}$$

If $|N| = 1$, we have

$$\delta\Lambda_{st}^{+000}K_s^{+000} + \delta\Lambda_{st}^{-000}K_s^{-000} = \mathrm{i}r^2\left(\delta\Lambda_{st}^{+000} - (-1)^{l+l'+s}\delta\Lambda_{st}^{-000}\right)\left[\Omega_0^l Z\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix} + \Omega_0^{l'} Z'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}\right], \tag{B43}$$

$$\delta\Lambda_{st}^{++-0}K_s^{++-0} + \delta\Lambda_{st}^{--+0}K_s^{--+0} = -\mathrm{i}r\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{++-0} - (-1)^{l+l'+s}\delta\Lambda_{st}^{--+0}\right)\left[\Omega_2^l(VZ' - WX')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(V'Z - W'X)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right],$$

$$\delta\Lambda_{st}^{+-+0}K_s^{+-+0} + \delta\Lambda_{st}^{-+-0}K_s^{-+-0} = -\mathrm{i}r^2\left(\delta\Lambda_{st}^{+-+0} - (-1)^{l+l'+s}\delta\Lambda_{st}^{-+-0}\right)\left[\Omega_0^{l'}FZ'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix} + \Omega_0^l F'Z\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]. \tag{B44}$$

For $|N| = 2$,

$$\delta\Lambda_{st}^{++00}K_s^{++00} + \delta\Lambda_{st}^{--00}K_s^{--00} = -\mathrm{i}r\left(\delta\Lambda_{st}^{++00} - (-1)^{l+l'+s}\delta\Lambda_{st}^{--00}\right)\left[\Omega_0^l\Omega_2^l W\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} - \Omega_0^{l'}\Omega_2^{l'}W'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{B45}$$

$$\delta\Lambda_{st}^{+0+0}K_s^{+0+0} + \delta\Lambda_{st}^{-0-0}K_s^{-0-0} = -\mathrm{i}r^2\Omega_0^l\Omega_0^{l'}(ZX' + XZ')\left(\delta\Lambda_{st}^{+0+0} - (-1)^{l+l'+s}\delta\Lambda_{st}^{-0-0}\right)\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}, \tag{B46}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-} + \delta\Lambda_{st}^{---+}K_s^{---+} = \mathrm{i}r\left(\delta\Lambda_{st}^{+++-} - (-1)^{l+l'+s}\delta\Lambda_{st}^{---+}\right)\left[\Omega_0^l\Omega_2^l WF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}\right.$$
$$\left. + \Omega_0^{l'}\Omega_2^{l'}\mathrm{i}W'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right]. \tag{B47}$$

For $|N| = 3$,

$$\delta\Lambda_{st}^{+++0}K_s^{+++0} + \delta\Lambda_{st}^{---0}K_s^{---0} = \mathrm{i}r\Omega_0^{l'}\Omega_0^l\left(\delta\Lambda_{st}^{+++0} - (-1)^{l+l'+s}\delta\Lambda_{st}^{---0}\right)\left[\Omega_2^l(WX' + V'Z)\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(W'X + VZ')\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right]. \tag{B48}$$

And for $|N| = 4$, we have

$$\delta\Lambda_{st}^{++++}K_s^{++++} + \delta\Lambda_{st}^{----}K_s^{----} = -\mathrm{i}(WV' + VW')\Omega_0^{l'}\Omega_2^{l'}\Omega_0^l\Omega_2^l\left(\delta\Lambda_{st}^{++++} - (-1)^{l+l'+s}\delta\Lambda_{st}^{----}\right)\begin{pmatrix} l & s & l' \\ -2 & 4 & -2\end{pmatrix}. \tag{B49}$$

APPENDIX C: PARTICULAR CASES

In the following subsections, we examinate different cases of coupling. We make use of the following relations (Montagner 1996; Montagner & Guillot 2000; Chen & Tromp 2007):

$$\delta\Lambda^{+000} + \delta\Lambda^{-000} = \mathrm{i}4(J_s - K_s)/\sqrt{2}, \tag{C1}$$

$$\delta\Lambda^{+000} - \delta\Lambda^{-000} = -4(J_c - K_c)/\sqrt{2}, \tag{C2}$$

$$\delta\Lambda^{++-0} + \delta\Lambda^{-+0} = \mathrm{i}M_s/\sqrt{2}, \tag{C3}$$

$$\delta\Lambda^{++-0} - \delta\Lambda^{-+0} = M_c/\sqrt{2}, \tag{C4}$$

$$\delta\Lambda^{+-+0} + \delta\Lambda^{-+-0} = -\mathrm{i}2(M_s + 2J_s)/\sqrt{2}, \tag{C5}$$

$$\delta\Lambda^{+-+0} - \delta\Lambda^{-+-0} = 2(2J_c - M_c)/\sqrt{2}, \tag{C6}$$

$$\delta\Lambda^{0+0+} + \delta\Lambda^{0-0-} = 2G_c, \tag{C7}$$

$$\delta\Lambda^{0+0+} - \delta\Lambda^{0-0-} = -\mathrm{i}2G_s, \tag{C8}$$

$$\delta\Lambda^{---+} + \delta\Lambda^{+++-} = -B_c, \tag{C9}$$

$$\delta\Lambda^{---+} - \delta\Lambda^{+++-} = -\mathrm{i}B_s, \tag{C10}$$

$$\delta\Lambda^{--00} + \delta\Lambda^{++00} = 2H_c, \tag{C11}$$

$$\delta\Lambda^{--00} - \delta\Lambda^{++00} = -\mathrm{i}2H_s, \tag{C12}$$

$$\delta\Lambda^{+++0} + \delta\Lambda^{---0} = \mathrm{i}4D_s/\sqrt{2}, \tag{C13}$$

$$\delta\Lambda^{+++0} - \delta\Lambda^{---0} = -4D_c/\sqrt{2}, \tag{C14}$$

$$\delta\Lambda^{++++} = \Lambda^{----} = 2E_c + 2\mathrm{i}E_s, \tag{C15}$$

and we employ Love's notation (Love 1927) to describe radial anisotropy. The Love coefficients A, C, N, L and F are related to the $\Lambda^{\alpha\beta\gamma\delta}$ through the following relationships (Mochizuki 1986).

$$\Lambda^{0000} = C, \tag{C16}$$

$$\Lambda^{++--} = 2N, \tag{C17}$$

$$\Lambda^{+-+-} = A - N, \tag{C18}$$

$$\Lambda^{+-00} = -F, \tag{C19}$$

$$\Lambda^{+0-0} = -L. \tag{C20}$$

C1 Like-type coupling with $l + l' + s$ even

Here, we make use of the symmetry properties of the 3j-Wigner symbols (e.g. Dahlen & Tromp 1998) and eq. (B3).

C1.1 Radial anisotropy

$$\delta\Lambda^{0000}_{st} K^{0000}_s(r) = r^2 \dot{U}\dot{U}' \begin{pmatrix} l & s & l' \\ 0 & 0 & 0 \end{pmatrix} \delta C_{st}, \tag{C21}$$

$$\delta\Lambda^{+-+-}_{st} K^{+-+-}_s(r) = r^2 F F' \begin{pmatrix} l & s & l' \\ 0 & 0 & 0 \end{pmatrix} (\delta A_{st} - \delta N_{st}), \tag{C22}$$

$$\delta\Lambda^{+-00}_{st} K^{+-00}_s(r) = r^2 (F\dot{U}' + \dot{U}F') \begin{pmatrix} l & s & l' \\ 0 & 0 & 0 \end{pmatrix} \delta F_{st}, \tag{C23}$$

$$\delta\Lambda^{++--}_{st} K^{++--}_s(r) = 2\Omega^l_0\Omega^l_2\Omega^{l'}_0\Omega^{l'}_2 (VV' + WW') \begin{pmatrix} l & s & l' \\ 2 & 0 & -2 \end{pmatrix} 2\delta N_{st}, \tag{C24}$$

$$\delta\Lambda_{st}^{+0-0}K_s^{+0-0}(r) = -\Omega_0^l\Omega_0^{l'}r^2(XX+ZZ')\begin{pmatrix} l & s & l' \\ -1 & 0 & 1\end{pmatrix}\delta L_{st}. \tag{C25}$$

We can rewrite those last equations to have an expression of the type $K_A(r)\delta A_{st} + K_C(r)\delta C_{st} + K_N(r)\delta N_{st} + K_L(r)\delta L_{st} + K_F(r)\delta F_{st}$. In that case, we have

$$K_A = r^2FF'\begin{pmatrix} l & s & l' \\ 0 & 0 & 0\end{pmatrix}, \tag{C26}$$

$$K_C = r^2\dot{U}\dot{U}'\begin{pmatrix} l & s & l' \\ 0 & 0 & 0\end{pmatrix}, \tag{C27}$$

$$K_N = -K_A + 4\Omega_0^l\Omega_2^l\Omega_0^{l'}\Omega_2^{l'}(VV'+WW')\begin{pmatrix} l & s & l' \\ 2 & 0 & -2\end{pmatrix}, \tag{C28}$$

$$K_L = \Omega_0^l\Omega_0^{l'}r^2(XX'+ZZ')\begin{pmatrix} l & s & l' \\ -1 & 0 & 1\end{pmatrix}, \tag{C29}$$

$$K_F = r^2(F\dot{U}'+F'\dot{U})\begin{pmatrix} l & s & l' \\ 0 & 0 & 0\end{pmatrix}. \tag{C30}$$

C1.2 Azimuthal anisotropy

$$K_s^{+000}\delta\Lambda_{st}^{+000} + K_s^{-000}\delta\Lambda_{st}^{-000} = -\left(\delta\Lambda_{st}^{+000}+\delta\Lambda_{st}^{-000}\right)r^2\left[\Omega_0^lX\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix} + \Omega_0^{l'}X'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}\right], \tag{C31}$$

$$K_s^{++-0}\delta\Lambda_{st}^{++-0} + K_s^{-+0}\delta\Lambda_{st}^{-+0} = -\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{++-0}+\delta\Lambda_{st}^{-+0}\right)r\left[\Omega_2^l(VX'+WZ')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(V'X+W'Z)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right], \tag{C32}$$

$$K_s^{+-+0}\delta\Lambda_{st}^{+-+0} + K_s^{-+-0}\delta\Lambda_{st}^{-+-0} = \left(\delta\Lambda_{st}^{+-+0}+\delta\Lambda_{st}^{-+-0}\right)r^2\left[\Omega_0^{l'}FX'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix} + \Omega_0^lF'X\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right], \tag{C33}$$

$$\delta\Lambda_{st}^{++00}K_s^{++00} + \delta\Lambda_{st}^{-00}K_s^{-00} = \left(\delta\Lambda_{st}^{++00}+\delta\Lambda_{st}^{-00}\right)r\left[\Omega_0^l\Omega_2^lV\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} + \Omega_2^{l'}\Omega_0^{l'}V'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C34}$$

$$\delta\Lambda_{st}^{0+0+}K_s^{0+0+} + \delta\Lambda_{st}^{0-0-}K_s^{0-0-} = \Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{0+0+}+\delta\Lambda_{st}^{0-0-}\right)r^2(XX'-ZZ')\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}, \tag{C35}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-} + \delta\Lambda_{st}^{---+}K_s^{---+} = -\left(\delta\Lambda_{st}^{+++-}+\delta\Lambda_{st}^{---+}\right)r\left[\Omega_0^l\Omega_2^lVF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} + \Omega_2^{l'}\Omega_0^{l'}V'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C36}$$

$$\delta\Lambda_{st}^{+++0}K_s^{+++0}+\delta\Lambda_{st}^{---0}K_s^{---0}=-\Omega_0^{l'}\Omega_l^0\left(\delta\Lambda_{st}^{+++0}+\delta\Lambda_{st}^{---0}\right)r\left[\Omega_2^l(VX'-WZ')\begin{pmatrix}l&s&l'\\-2&3&-1\end{pmatrix}\right.$$
$$\left.+\Omega_2^{l'}(V'X-W'Z)\begin{pmatrix}l&s&l'\\-1&3&-2\end{pmatrix}\right], \tag{C37}$$

$$\delta\Lambda_{st}^{++++}K_s^{++++}+\delta\Lambda_{st}^{----}K_s^{----}=2\Omega_2^{l'}\Omega_0^{l'}\Omega_2^l\Omega_l^0(\delta\Lambda_{st}^{++++}+\delta\Lambda_{st}^{----})(VV'-WW')\begin{pmatrix}l&s&l'\\-2&4&-2\end{pmatrix}. \tag{C38}$$

These equations can be written in terms of elastic parameters $\mathrm{i}J_s$, $\mathrm{i}M_s$, $\mathrm{i}D_s$, $\mathrm{i}K_s$, G_c, B_c, H_c and $E_c+\mathrm{i}E_s$

Elastic coefficient	Kernel
$\mathrm{i}\,(K_s)_{st}$	$\frac{4}{\sqrt{2}}r^2\left[\Omega_0^l X\dot{U}'\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}+\Omega_0^{l'}X'\dot{U}\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}\right]$
$\mathrm{i}\,(M_s)_{st}$	$-\frac{1}{\sqrt{2}}\Omega_0^l\Omega_0^{l'}r\left[\Omega_2^l(VX'+WZ')\begin{pmatrix}l&s&l'\\-2&1&1\end{pmatrix}+\Omega_2^{l'}(V'X+W'Z)\begin{pmatrix}l&s&l'\\1&1&-2\end{pmatrix}\right]$
	$-\frac{2}{\sqrt{2}}r^2\left[\Omega_0^{l'}FX'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^lF'X\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right]$
$\mathrm{i}\,(J_s)_{st}$	$-\frac{4}{\sqrt{2}}r^2\left[\Omega_0^{l'}(F+\dot{U})X'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^l(F'+\dot{U}')X\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right]$
$(H_c)_{st}$	$2\Omega_0^l\Omega_2^lrV\dot{U}'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}+2\Omega_2^{l'}\Omega_0^{l'}rV'\dot{U}\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}$
$(G_c)_{st}$	$2\Omega_0^l\Omega_0^{l'}r^2(XX'-ZZ')\begin{pmatrix}l&s&l'\\-1&2&-1\end{pmatrix}$
$(B_c)_{st}$	$\Omega_0^l\Omega_2^lrVF'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}+\Omega_2^{l'}\Omega_0^{l'}rV'F\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}$
$\mathrm{i}\,(D_s)_{st}$	$-\frac{4}{\sqrt{2}}\Omega_0^{l'}\Omega_0^lr\left[\Omega_2^l(VX'-WZ')\begin{pmatrix}l&s&l'\\-2&3&-1\end{pmatrix}+\Omega_2^{l'}(V'X-W'Z)\begin{pmatrix}l&s&l'\\-1&3&-2\end{pmatrix}\right]$
$(E_c)_{st}+\mathrm{i}\,(E_s)_{st}$	$4\Omega_2^{l'}\Omega_0^{l'}\Omega_2^l\Omega_0^l(VV'-WW')\begin{pmatrix}l&s&l'\\-2&4&-2\end{pmatrix}$

C2 Like-type coupling with $l+l'+s$ odd

When $l+l'+s$ takes odd values, the sensitivity kernels for radial anisotropy are zero. Like-type coupled modes are therefore not sensitive to radial anisotropy when $l+l'+s$ is odd, and we are left with equations for azimuthal anisotropy only.

$$K_s^{+000}\Lambda_{st}^{+000}+K_s^{-000}\Lambda_{st}^{-000}=-\left(\Lambda_{st}^{+000}-\Lambda_{st}^{-000}\right)r^2\left[\Omega_0^lX\dot{U}'\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}+\Omega_0^{l'}X'\dot{U}\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}\right], \tag{C39}$$

$$K_s^{++-0}\Lambda_{st}^{++-0}+K_s^{-+0}\Lambda_{st}^{-+0}=-\Omega_0^l\Omega_0^{l'}\left(\Lambda_{st}^{++-0}-\Lambda_{st}^{-+0}\right)r\left[\Omega_2^l(VX'+WZ')\begin{pmatrix}l&s&l'\\-2&1&1\end{pmatrix}+\Omega_2^{l'}(V'X+W'Z)\begin{pmatrix}l&s&l'\\1&1&-2\end{pmatrix}\right], \tag{C40}$$

$$K_s^{+-+0}(r)\Lambda_{st}^{+-+0}+K_s^{-+-0}(r)\Lambda_{st}^{-+-0}=\left(\Lambda_{st}^{+-+0}-\Lambda_{st}^{-+-0}\right)r^2\left[\Omega_0^{l'}FX'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^lF'X\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right], \tag{C41}$$

$$\delta\Lambda_{st}^{++00}K_s^{++00}+\delta\Lambda_{st}^{-00}K_s^{-00}=-\left(\delta\Lambda_{st}^{++00}-\delta\Lambda_{st}^{-00}\right)r\left[\Omega_0^l\Omega_2^l V\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}-\Omega_0^{l'}\Omega_2^{l'}V'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C42}$$

$$\delta\Lambda_{st}^{0+0+}K_s^{0+0+}+\delta\Lambda_{st}^{0-0-}K_s^{0-0-}=\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{0+0+}-\delta\Lambda_{st}^{0-0-}\right)r^2(XX'-ZZ')\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}, \tag{C43}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-}+\delta\Lambda_{st}^{---+}K_s^{---+}=-\left(\delta\Lambda_{st}^{+++-}-\delta\Lambda_{st}^{---+}\right)r\left[\Omega_0^l\Omega_2^l VF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}+\Omega_0^{l'}\Omega_2^{l'}V'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C44}$$

$$\begin{aligned}\delta\Lambda_{st}^{+++0}K_s^{+++0}+\delta\Lambda_{st}^{---0}K_s^{---0}=&-\Omega_0^{l'}\Omega_l^0\left(\delta\Lambda_{st}^{+++0}-\delta\Lambda_{st}^{---0}\right)r\left[\Omega_2^l(VX'-WZ')\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}\right.\\&\left.+\Omega_2^{l'}(V'X-W'Z)\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right],\end{aligned} \tag{C45}$$

$$\delta\Lambda_{st}^{----}K_s^{----}+\delta\Lambda_{st}^{++++}K_s^{++++}=0. \tag{C46}$$

These equations can be written in terms of elastic parameters J_c, M_c, D_c, K_c, $\mathrm{i}G_s$, $\mathrm{i}B_s$ and $\mathrm{i}H_s$

Elastic coefficient	Kernel
$(K_c)_{st}$	$-\frac{4}{\sqrt{2}}r^2\left[\Omega_0^l X\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}+\Omega_0^{l'}X'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}\right]$
$(M_c)_{st}$	$-\frac{1}{\sqrt{2}}\Omega_0^l\Omega_0^{l'}r\left[\Omega_2^l(VX'+WZ')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}+\Omega_2^{l'}(V'X+W'Z)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right]$
	$-\frac{2}{\sqrt{2}}r^2\left[\Omega_0^{l'}FX'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}+\Omega_0^lF'X\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]$
$(J_c)_{st}$	$\frac{4}{\sqrt{2}}r^2\left[\Omega_0^{l'}\left(F+\dot{U}\right)X'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}+\Omega_0^l\left(F'+\dot{U}'\right)X\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]$
$\mathrm{i}\,(H_s)_{st}$	$-2\Omega_0^l\Omega_2^l rV\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}-2\Omega_0^{l'}\Omega_2^{l'}rV'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}$
$\mathrm{i}\,(G_s)_{st}$	$-2\Omega_0^l\Omega_0^{l'}r^2(XX'-ZZ')\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}$
$\mathrm{i}\,(B_s)_{st}$	$-\Omega_0^l\Omega_2^l rVF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}-\Omega_0^{l'}\Omega_2^{l'}rV'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}$
$(D_c)_{st}$	$\Omega_0^{l'}\Omega_l^0\frac{4}{\sqrt{2}}r\left[\Omega_2^l(VX'-WZ')\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}+\Omega_2^{l'}(V'X-W'Z)\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right]$

C3 Cross-type coupling with $l + l' + s$ even

In this case, coupled modes are not sensitive to radial anisotropy. The sensitivity to azimuthal anisotropy is described by the following equations:

$$\delta\Lambda_{st}^{+000}K_s^{+000}+\delta\Lambda_{st}^{-000}K_s^{-000},=\mathrm{i}\left(\delta\Lambda_{st}^{+000}-\delta\Lambda_{st}^{-000}\right)r^2\left[\Omega_0^lZ\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}+\Omega_0^{l'}Z'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & -1 & 1\end{pmatrix}\right], \tag{C47}$$

$$\delta\Lambda_{st}^{++-0}K_s^{++-0}+\delta\Lambda_{st}^{-+0}K_s^{-+0}=-\mathrm{i}\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{++-0}-\delta\Lambda_{st}^{-+0}\right)r\left[\Omega_2^l(VZ'-WX')\begin{pmatrix}l&s&l'\\-2&1&1\end{pmatrix}\right.$$
$$\left.+\Omega_2^{l'}(V'Z-W'X)\begin{pmatrix}l&s&l'\\1&1&-2\end{pmatrix}\right], \tag{C48}$$

$$\delta\Lambda_{st}^{+-+0}K_s^{+-+0}+\delta\Lambda_{st}^{-+-0}K_s^{-+-0}=-\mathrm{i}\left(\delta\Lambda_{st}^{+-+0}-\delta\Lambda_{st}^{-+-0}\right)r^2\left[\Omega_0^{l'}FZ'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^lF'Z\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right], \tag{C49}$$

$$\delta\Lambda_{st}^{++00}K_s^{++00}+\delta\Lambda_{st}^{-00}K_s^{-00}=-\mathrm{i}\left(\delta\Lambda_{st}^{++00}-\delta\Lambda_{st}^{-00}\right)r\left[\Omega_0^l\Omega_2^lW\dot{U}'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}-\Omega_0^{l'}\Omega_2^{l'}W'\dot{U}\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}\right], \tag{C50}$$

$$\delta\Lambda_{st}^{+0+0}K_s^{+0+0}+\delta\Lambda_{st}^{-0-0}K_s^{-0-0}=-\mathrm{i}\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{+0+0}-\delta\Lambda_{st}^{-0-0}\right)r^2(ZX'+XZ')\begin{pmatrix}l&s&l'\\-1&2&-1\end{pmatrix}, \tag{C51}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-}+\delta\Lambda_{st}^{---+}K_s^{---+}=\mathrm{i}\left(\delta\Lambda_{st}^{+++-}-\delta\Lambda_{st}^{---+}\right)r\left[\Omega_0^l\Omega_2^l\mathrm{i}WF'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}+\Omega_0^{l'}\Omega_2^{l'}\mathrm{i}W'F\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}\right], \tag{C52}$$

$$\delta\Lambda_{st}^{+++0}K_s^{+++0}+\delta\Lambda_{st}^{---0}K_s^{---0}=\mathrm{i}\Omega_0^{l'}\Omega_0^l\left(\delta\Lambda_{st}^{+++0}-\delta\Lambda_{st}^{---0}\right)r\left[\Omega_2^l(WX'+V'Z)\begin{pmatrix}l&s&l'\\-2&3&-1\end{pmatrix}\right.$$
$$\left.+\Omega_2^{l'}(W'X+VZ')\begin{pmatrix}l&s&l'\\-1&3&-2\end{pmatrix}\right], \tag{C53}$$

$$\delta\Lambda_{st}^{++++}K_s^{++++}+\delta\Lambda_{st}^{----}K_s^{----}=-\mathrm{i}(WV'+VW')\Omega_0^{l'}\Omega_2^{l'}\Omega_0^l\Omega_2^l\left(\delta\Lambda_{st}^{++++}-\delta\Lambda_{st}^{----}\right)\begin{pmatrix}l&s&l'\\-2&4&-2\end{pmatrix}=0. \tag{C54}$$

In terms of elastic parameters $\mathrm{i}J_c$, $\mathrm{i}M_c$, $\mathrm{i}D_c$, $\mathrm{i}K_c$, G_s, B_s and H_s, these equations become

Elastic coefficient	Kernel
$\mathrm{i}\,(K_c)_{st}$	$\frac{4}{\sqrt{2}}r^2\left[\Omega_0^lZ\dot{U}'\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}+\Omega_0^{l'}Z'\dot{U}\begin{pmatrix}l&s&l'\\0&-1&1\end{pmatrix}\right]$
$\mathrm{i}\,(M_c)_{st}$	$-\frac{1}{\sqrt{2}}\Omega_0^l\Omega_0^{l'}r\left[\Omega_2^l(VZ'-WX')\begin{pmatrix}l&s&l'\\-2&1&1\end{pmatrix}+\Omega_2^{l'}(V'Z-W'X)\begin{pmatrix}l&s&l'\\1&1&-2\end{pmatrix}\right]$
	$+\frac{2}{\sqrt{2}}r^2\left[\Omega_0^{l'}FZ'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^lF'Z\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right]$
$\mathrm{i}\,(J_c)_{st}$	$-\frac{4}{\sqrt{2}}r^2\left[\Omega_0^{l'}\left(F+\dot{U}\right)Z'\begin{pmatrix}l&s&l'\\0&1&-1\end{pmatrix}+\Omega_0^l\left(F'+\dot{U}'\right)Z\begin{pmatrix}l&s&l'\\-1&1&0\end{pmatrix}\right]$
$(H_s)_{st}$	$-2\Omega_0^l\Omega_2^lrW\dot{U}'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}-2\Omega_0^{l'}\Omega_2^{l'}rW'\dot{U}\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}$
$(G_s)_{st}$	$-2\Omega_0^l\Omega_0^{l'}r^2(ZX'+XZ')\begin{pmatrix}l&s&l'\\-1&2&-1\end{pmatrix}$
$(B_s)_{st}$	$-\Omega_0^l\Omega_2^lrWF'\begin{pmatrix}l&s&l'\\-2&2&0\end{pmatrix}-\Omega_0^{l'}\Omega_2^{l'}rW'F\begin{pmatrix}l&s&l'\\0&2&-2\end{pmatrix}$
$\mathrm{i}\,(D_c)\,st$	$4\sqrt{2}\Omega_0^{l'}\Omega_0^lr\left[\Omega_2^l(WX'+V'Z)\begin{pmatrix}l&s&l'\\-2&3&-1\end{pmatrix}+\Omega_2^{l'}(W'X+VZ')\begin{pmatrix}l&s&l'\\-1&3&-2\end{pmatrix}\right]$

C4 Cross-type coupling with $l + l' + s$ odd

C4.1 Radial anisotropy

Eqs (B38)–(B42) become

$$\delta\Lambda_{st}^{0000}K_s^{0000} = 0, \tag{C55}$$

$$\delta\Lambda_{st}^{+-00}K_s^{+-00} = 0, \tag{C56}$$

$$\delta\Lambda_{st}^{+-+-}K_s^{+-+-} = 0, \tag{C57}$$

$$\delta\Lambda_{st}^{++-}K_s^{++-} = \mathrm{i}(VW' - V'W)B_{lsl'}^{2-}\delta N_{st}, \tag{C58}$$

$$\delta\Lambda_{st}^{+0-0}K_s^{+0-0} = \mathrm{i}\delta L_{st}r^2(XZ' - X'Z)B_{lsl'}^{1-}. \tag{C59}$$

Therefore,

$$K_A = 0, \tag{C60}$$

$$K_C = 0, \tag{C61}$$

$$K_N = \mathrm{i}(VW' - WV')B_{lsl'}^{2-}, \tag{C62}$$

$$K_L = \mathrm{i}r^2(XZ' - ZX')B_{lsl'}^{1-}, \tag{C63}$$

$$K_F = 0. \tag{C64}$$

This shows that the coupling between spheroidal and toroidal modes are sensitive to shear wave anisotropy but not to *P*-wave anisotropy, when $l + l' + s$ is odd.

C4.2 Azimuthal anisotropy

$$\delta\Lambda_{st}^{+000}K_s^{+000} + \delta\Lambda_{st}^{-000}K_s^{-000} = \mathrm{i}\left(\delta\Lambda_{st}^{+000} + \delta\Lambda_{st}^{-000}\right)r^2\left[\Omega_0^l Z\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix} + \Omega_0^{l'}Z'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & -1 & 1\end{pmatrix}\right], \tag{C65}$$

$$\delta\Lambda_{st}^{++-0}K_s^{++-0} + \delta\Lambda_{st}^{-+0}K_s^{-+0} = -\mathrm{i}\Omega_0^l\Omega_0^{l'}\left(\delta\Lambda_{st}^{++-0} + \delta\Lambda_{st}^{-+0}\right)r\left[\Omega_2^l(VZ' - WX')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}\right.$$
$$\left. + \Omega_2^{l'}(V'Z - W'X)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right], \tag{C66}$$

$$\delta\Lambda_{st}^{+-+0}K_s^{+-+0} + \delta\Lambda_{st}^{-+-0}K_s^{-+-0} = -\mathrm{i}(\delta\Lambda_{st}^{+-+0} + \delta\Lambda_{st}^{-+-0})r^2\left[\Omega_0^{l'}FZ'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix} + \Omega_0^l F'Z\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right], \tag{C67}$$

$$\delta\Lambda_{st}^{++00}K_s^{++00} + \delta\Lambda_{st}^{-00}K_s^{-00} = -\mathrm{i}\left(\delta\Lambda_{st}^{++00} + \delta\Lambda_{st}^{-00}\right)r\left[\Omega_0^l\Omega_2^l W\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} - \Omega_0^{l'}\Omega_2^{l'}W'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C68}$$

$$\delta\Lambda_{st}^{+0+0}K_s^{+0+0} + \delta\Lambda_{st}^{-0-0}K_s^{-0-0} = -\mathrm{i}\Omega_0^{l'}\Omega_0^l\left(\delta\Lambda_{st}^{+0+0} + \delta\Lambda_{st}^{-0-0}\right)r^2(X'Z + XZ')\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}, \tag{C69}$$

$$\delta\Lambda_{st}^{+++-}K_s^{+++-} + \delta\Lambda_{st}^{---+}K_s^{---+} = \mathrm{i}\left(\delta\Lambda_{st}^{+++-} + \delta\Lambda_{st}^{---+}\right)r\left[\Omega_0^l\Omega_2^l WF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix} + \Omega_0^{l'}\Omega_2^{l'}W'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}\right], \tag{C70}$$

$$\delta\Lambda_{st}^{+++0}K_s^{+++0}+\delta\Lambda_{st}^{---0}K_s^{---0}=\mathrm{i}\Omega_0^{l'}\Omega_0^{l}\left(\delta\Lambda_{st}^{+++0}+\delta\Lambda_{st}^{---0}\right)r\left[\Omega_2^{l}(WX'+V'Z)\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}\right.$$
$$\left.+\Omega_2^{l'}(W'X+VZ')\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right], \tag{C71}$$

$$\delta\Lambda_{st}^{++++}K_s^{++++}+\delta\Lambda_{st}^{----}K_s^{----}=-\mathrm{i}\Omega_0^{l'}\Omega_2^{l'}\Omega_0^{l}\Omega_2^{l}\left(\delta\Lambda_{st}^{++++}+\delta\Lambda_{st}^{----}\right)(WV'+W'V)\begin{pmatrix} l & s & l' \\ -2 & 4 & -2\end{pmatrix}. \tag{C72}$$

In terms of elastic parameters J_S, M_s, D_s, K_s, $\mathrm{i}G_c$, $\mathrm{i}B_c$, $\mathrm{i}H_c$ and $\mathrm{i}\,E_c-E_s$, these equations become

Elastic coefficient	kernel
$(K_s)_{st}$	$\frac{4}{\sqrt{2}}r^2\left[\Omega_0^{l}Z\dot{U}'\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}+\Omega_0^{l'}Z'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & -1 & 1\end{pmatrix}\right]$
$(M_s)_{st}$	$\frac{1}{\sqrt{2}}\Omega_0^{l}\Omega_0^{l'}r\left[\Omega_2^{l}(VZ'-WX')\begin{pmatrix} l & s & l' \\ -2 & 1 & 1\end{pmatrix}+\Omega_2^{l'}(V'Z-W'X)\begin{pmatrix} l & s & l' \\ 1 & 1 & -2\end{pmatrix}\right]$
	$-\frac{2}{\sqrt{2}}r^2\left[\Omega_0^{l'}FZ'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}+\Omega_0^{l}F'Z\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]$
$(J_s)_{st}$	$-\frac{4}{\sqrt{2}}r^2\left[\Omega_0^{l'}\left(F+\dot{U}\right)Z'\begin{pmatrix} l & s & l' \\ 0 & 1 & -1\end{pmatrix}+\Omega_0^{l}\left(F'+\dot{U}'\right)Z\begin{pmatrix} l & s & l' \\ -1 & 1 & 0\end{pmatrix}\right]$
$\mathrm{i}\,(H_c)_{st}$	$-2\Omega_0^{l}\Omega_2^{l}rW\dot{U}'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}-2\Omega_0^{l'}\Omega_2^{l'}rW'\dot{U}\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}$
$\mathrm{i}\,(G_c)_{st}$	$-2\Omega_0^{l'}\Omega_0^{l}r^2(X'Z+XZ')\begin{pmatrix} l & s & l' \\ -1 & 2 & -1\end{pmatrix}$
$\mathrm{i}\,(B_c)_{st}$	$-\Omega_0^{l}\Omega_2^{l}rWF'\begin{pmatrix} l & s & l' \\ -2 & 2 & 0\end{pmatrix}-\Omega_0^{l'}\Omega_2^{l'}rW'F\begin{pmatrix} l & s & l' \\ 0 & 2 & -2\end{pmatrix}$
$(D_s)_{st}$	$-\frac{4}{\sqrt{2}}\Omega_0^{l'}\Omega_0^{l}r\left[\Omega_2^{l}(WX'+V'Z)\begin{pmatrix} l & s & l' \\ -2 & 3 & -1\end{pmatrix}+\Omega_2^{l'}(W'X+VZ')\begin{pmatrix} l & s & l' \\ -1 & 3 & -2\end{pmatrix}\right]$
$\mathrm{i}\,(E_c)_{st}-(E_s)_{st}$	$-4\Omega_0^{l'}\Omega_2^{l'}\Omega_0^{l}\Omega_2^{l}(WV'+W'V)\begin{pmatrix} l & s & l' \\ -2 & 4 & -2\end{pmatrix}$