



**HAL**  
open science

## Residual homogenization for seismic forward and inverse problems in layered media

Yann Capdeville, Eléonore Stutzmann, Nian Wang, Jean-Paul Montagner

### ► To cite this version:

Yann Capdeville, Eléonore Stutzmann, Nian Wang, Jean-Paul Montagner. Residual homogenization for seismic forward and inverse problems in layered media. *Geophysical Journal International*, 2013, 194 (1), pp.470-487. 10.1093/gji/ggt102 . insu-01396948

**HAL Id: insu-01396948**

**<https://insu.hal.science/insu-01396948>**

Submitted on 15 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Residual homogenization for seismic forward and inverse problems in layered media

Yann Capdeville,<sup>1</sup> Éléonore Stutzmann,<sup>2</sup> Nian Wang<sup>2</sup> and Jean-Paul Montagner<sup>2</sup>

<sup>1</sup>Laboratoire de Planétologie et Géodynamique de Nantes, CNRS, Université de Nantes, UMR-6112, France. E-mail: Yann.Capdeville@univ-nantes.fr

<sup>2</sup>Équipe de Sismologie, Institut de Physique du Globe de Paris, CNRS, UMR-7154, France

Accepted 2013 March 14. Received 2013 March 4; in original form 2013 January 29

### SUMMARY

An elastic wavefield propagating in an inhomogeneous elastic medium is only sensitive in an effective way to inhomogeneities much smaller than its minimum wavelength. The corresponding effective medium, or homogenized medium, can be computed thanks to the non-periodic homogenization technique. In the seismic imaging context, limiting ourselves to layered media, we numerically show that a tomographic elastic model which results of the inversion of limited frequency band seismic data is an homogenized model. Moreover, we show that this homogenized model is the same as the model that can be computed with the non-periodic homogenization technique. We first introduce the notion of residual homogenization, which is computing the effective properties of the difference between a reference model and a ‘real’ model. This is necessary because most imaging technique parametrizations use a reference model that often contains small scales, such as elastic discontinuities. We then use a full-waveform inversion method to numerically show that the result of the inversion is indeed the homogenized residual model. The full-waveform inversion method used here has been specifically developed for that purpose. It is based on the iterative Gauss–Newton least-square non-linear optimization technique, using full normal mode coupling to compute the partial Hessian and gradient. The parametrization has been designed according to the residual homogenized parameters allowing to build a real multiscale inversion with progressive frequency band enrichment along the Gauss–Newton iterations.

**Key words:** Numerical solutions; Inverse theory; Seismic tomography; Computational seismology; Theoretical Seismology; Wave propagation.

### 1 INTRODUCTION

Since Backus (1962)’s work, it is well known that elastic waves of a given minimum wavelength  $\lambda_{\min}$  propagating in the Earth are sensitive to inhomogeneities of scale much smaller than  $\lambda_{\min}$  only in an effective way. For a given fine-scale layered medium, far away from the free surface and from the source, Backus (1962) showed how to compute the corresponding effective medium. Backus’s result can be extended to more general cases thanks to the two-scale homogenization technique for non-periodic media (Capdeville & Marigo 2007; Capdeville *et al.* 2010a,b; Guillot *et al.* 2010), and it is now possible, for a given general elastic 2-D or 3-D medium and a given minimum wavelength, to compute the corresponding 2-D or 3-D effective (or homogenized) medium for a given  $\lambda_{\min}$ . To solve the seismic forward problem of the elastic wave equation, the knowledge of the effective model is an important advantage. Indeed, the homogenized model is smooth and does not contain any elastic discontinuity, which releases drastically the meshing constraint and reduces the computational cost (Capdeville *et al.*

2010b). For the inverse problem, homogenization allows to mix results obtained at different scales (Fichtner *et al.* 2013). An intuitive interpretation of the homogenization results is that the effective medium computed according to  $\lambda_{\min}$  is the medium ‘seen’ by the wavefield of minimum wavelength  $\lambda_{\min}$ . This interpretation has potentially important consequences for the seismic imaging inverse problems. Indeed, it is not possible, using seismic records only, to have better information about the elastic model than what is ‘seen’ by the wavefield. As a consequence, a seismic inversion like a full-waveform tomographic method, can retrieve at best an effective medium but has no access to smaller scales. Moreover, this effective medium can be computed with the direct homogenization technique from the fine scale model. If it is difficult to mathematically prove this conjecture, it is possible to numerically show it is indeed true. The main objective of the present work is to numerically show that, knowing a fine scale model, computing the homogenized medium gives the same result as inverting synthetic seismic data (computed in the fine scale model) for the same minimum wavelength.

As a first step towards proving that an inverted elastic model obtained from a tomographic method is indeed an homogenized model in the general case, we limit ourselves in this paper to the layered media case. We work at the global Earth scale, that is we consider spherically symmetric earth models, knowing that our results will be also valid for layered medium. The spherically symmetric case has many advantages: the direct homogenization is quite simple and the full-waveform inverse problem can be solved efficiently with a limited amount of computing resources. Working with spherically symmetric global earth models is also relevant to many current tomography techniques used in global seismology. Indeed, many source–receiver ‘path average’ inversion methods are based on the assumption that an average 1-D model can be used between each source–receiver pair. For such methods, typical measurements are dispersion curves of surface wave fundamental and higher modes, traveltimes of long-period body waves and eigenfrequencies (e.g. Cara *et al.* 1980; Nakanishi & Anderson 1982; Cara & L ev eque 1987; Montagner & Tanimoto 1990; Nolet 1990; Stutzmann & Montagner 1993; Trampert & Woodhouse 1995; Ekstr om *et al.* 1997; van Heijst & Woodhouse 1997; Yoshizawa & Kennett 2002; Beucler *et al.* 2003; Lebedev & van der Hilst 2008). More sophisticated methods, at global and exploration geophysics scales, based on 2-D or 3-D kernels and full-waveform methods (e.g. Li & Romanowicz 1996; Pratt *et al.* 1998; Montelli *et al.* 2004; Capdeville *et al.* 2005; Tromp *et al.* 2005; Fichtner *et al.* 2009; Virieux & Operto 2009; Leki c & Romanowicz 2011) are more and more taking over simple path average methods. Nevertheless, path average methods still have an important future thanks to tomographic methods based on noise cross-correlation measurements (e.g. Shapiro *et al.* 2005; Nishida *et al.* 2009; Schimmel *et al.* 2011). This is mainly due to the fact that the noise cross-correlation waveforms are difficult to use because the Earth’s noise source distribution is not perfect and poorly known (Cupillard 2008). For all the classical ‘path average’ tomographic methods, a reference model is necessary and quite important to design the radial parametrization. What is inverted for is the difference, or residual, between the real earth model and the reference model. Therefore, the inversion is expected to retrieve at best what the wavefield ‘sees’ of the difference between the real earth and the reference model. This leads us to introduce the new concept of residual homogenization, that is the homogenization of the difference between two models. We first introduce this new concept for the forward homogenization problem case, and then we introduce it for a full-waveform inverse problem.

The paper is organized as follows: in Section 2, the forward method designed to build a residual homogenized model for a given reference model is introduced and tested. In Section 3, the full-waveform inversion is introduced. The inversion is tested for two models that include small-scale inhomogeneities and the resulting inverted models are compared with the homogenized model in the same frequency band, showing that they are the same up to a given accuracy. In Section 4 are shown the residual homogenization effects on a gallery of earth models.

## 2 REDISUAL HOMOGENIZATION FOR THE FORWARD PROBLEM IN LAYERED MEDIA

### 2.1 Equations for elastic waves in layered media

In this section, we first recall the first-order system formulation of the elastic wave equations as it is classically done for the normal mode theory. When gravity and anelasticity are not taken into ac-

count, the wave equation in an elastic domain  $\Omega$  of boundary  $\partial\Omega$  can be written as

$$\rho \ddot{\mathbf{u}} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{f}, \quad (1)$$

$$\boldsymbol{\sigma} = \mathbf{c} : \boldsymbol{\epsilon}(\mathbf{u}), \quad (2)$$

where  $\rho$  is the density,  $\mathbf{u}$  the displacement field,  $\ddot{\mathbf{u}}$  the acceleration field,  $\boldsymbol{\sigma}$  the stress tensor,  $\mathbf{f}$  the source force,  $\mathbf{c}$  the fourth-order elastic tensor, ‘:’ the double indices contraction ( $\sigma_{ij} = \sum_{kl} c_{ijkl} \epsilon_{kl}$ ) and  $\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + {}^T \nabla \mathbf{u})$  the strain tensor with  ${}^T$  the transpose operator. We impose a free surface boundary condition on  $\partial\Omega$ ,  $\boldsymbol{\sigma} \cdot \mathbf{n} = 0$ , where  $\mathbf{n}$  is the normal vector to  $\partial\Omega$ . We assume that  $\mathbf{f}$  depends upon both time and space but we assume that the density and elastic properties are not time-dependent.

In this paper, similarly to Capdeville & Marigo (2007), we limit our work to layered media, in other words to 1-D media. All the examples and validation tests presented in this paper are performed in spherically symmetric global earth models of radius  $r_\Omega$ . We therefore use a spherical coordinates system  $\mathbf{r} = (r, \theta, \phi)$  where  $r$  is the radius,  $\theta$  the colatitude and  $\phi$  the longitude. This is absolutely not a limitation and all results presented here can be applied without modification to Cartesian or cylindrical coordinate systems for other types of layered models. In spherically symmetric layered media, we have  $\mathbf{c}(\mathbf{r}) = \mathbf{c}(r)$  and  $\rho(\mathbf{r}) = \rho(r)$ .

In such a framework, the classical wave equation can be rewritten as a first-order system of equations. To do so, the solution to (1) and (2) with free surface conditions is often sought in the frequency and the spectral domains for the horizontal directions (e.g. Takeuchi & Saito 1972). In spherically symmetric models, we use

$$\mathbf{u}(\mathbf{r}, \omega) =$$

$$\left[ U_l^m(r, \omega) \mathbf{e}_r + V_l^m(r, \omega) \nabla_1 - W_l^m(r, \omega) (\mathbf{e}_r \times \nabla_1) \right] Y_l^m(\theta, \phi), \quad (3)$$

where  $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi)$  is the spherical coordinate unit vector set,  $\nabla_1$  is the gradient operator on the unit sphere,  $Y_l^m$  the spherical harmonic of angular order  $l$  and azimuthal order  $m$  (e.g. Dahlen & Tromp 1998).  $\frac{\sqrt{l(l+1)}}{r}$  can be seen as the horizontal wavenumber. The radial traction  $\mathbf{T} = \boldsymbol{\sigma} \cdot \mathbf{e}_r$  can also be written under the form

$$\mathbf{T}(\mathbf{r}, \omega) =$$

$$\left[ T_{U_l}^m(r, \omega) \mathbf{e}_r + T_{V_l}^m(r, \omega) \nabla_1 - T_{W_l}^m(r, \omega) (\mathbf{e}_r \times \nabla_1) \right] Y_l^m(\theta, \phi). \quad (4)$$

Using (3) and (4) into (1) and (2), we obtain, in the frequency domain, for each  $l$ , two independent systems of equations, one for  $(U_l, T_{U_l}, V_l, T_{V_l})$  (spheroidal case) and one for  $(W_l, T_{W_l})$  (toroidal case), independent of  $m$ , that can be rewritten as a first-order system of equations

$$\frac{\partial_q \mathcal{Y}_l}{\partial r}(r, \omega) = {}_q \mathbf{A}_l(r, \omega) {}_q \mathcal{Y}_l(r, \omega), \quad (5)$$

where  $q$  can take two values,  $s$  for the spheroidal problem,  $t$  for the toroidal problem. We have

$${}_s \mathcal{Y}_l = {}^T (r U_l, r T_{U_l}, r V_l, r T_{V_l}) \quad (6)$$

for the spheroidal case and

$${}_t \mathcal{Y}_l = {}^T (r W_l, r T_{W_l}) \quad (7)$$

for the toroidal case, with  $\gamma_l = \sqrt{l(l+1)}$ . Expressions for  ${}_q \mathbf{A}_l$  matrices can, for example, be found in Takeuchi & Saito (1972), Aki & Richards (1980) or in Appendix A. They depend only on the

radius, on a non-linear combination of the five elastic parameters  $A$ ,  $C$ ,  $F$ ,  $L$ ,  $N$  and on the density. These elastic parameters, necessary to describe a vertical transversely isotropic (VTI) media, can be linked to the wave speeds in the medium by

$$V_{pv} = \sqrt{C/\rho} \text{ for vertically travelling } P \text{ waves,}$$

$$V_{sv} = \sqrt{L/\rho} \text{ for vertically polarized } S \text{ waves,}$$

$$V_{ph} = \sqrt{A/\rho} \text{ for horizontally travelling } P \text{ waves,}$$

$$V_{sh} = \sqrt{N/\rho} \text{ for horizontally polarized } S \text{ waves} \quad (8)$$

and to

$$\eta = \frac{F}{A - 2L}. \quad (9)$$

In the isotropic case, we have  $A = C = \lambda + 2\mu$ ,  $L = N = \mu$  and  $\eta = 1$ , where  $\lambda$  and  $\mu$  are the Lamé elastic coefficients. From the form of the matrices  ${}_q\mathbf{A}_l$  given in Appendix A, it can be seen that the problem linearly depends upon the following six parameters:

$$p_1(r) = \rho(r), \quad (10)$$

$$p_2(r) = \frac{1}{C}(r), \quad (11)$$

$$p_3(r) = \frac{1}{L}(r), \quad (12)$$

$$p_4(r) = \left( A - \frac{F^2}{C} \right) (r), \quad (13)$$

$$p_5(r) = \frac{F}{C}(r), \quad (14)$$

$$p_6(r) = N(r), \quad (15)$$

which define the parameter vector  $\mathbf{p}(r)$ . The six components of  $\mathbf{p}(r)$  are the same as the one found by Backus (1962) considering layered media and long waves. For that reason,  $p_i$  are here called the Backus parameters. As it can be seen from eqs (10)–(15), parameters  $p_i$  non-linearly depend upon the  $A$ ,  $C$ ,  $F$ ,  $L$  and  $N$  elastic parameters and on the density  $\rho$ . The vector  $\mathbf{p}$  fully defines the earth model and the operators  ${}_q\mathbf{A}_l$  and vice versa.

The solutions to (5) must be regular at  $r = 0$  and the free surface boundary conditions impose that radial traction must vanish for  $r = r_\Omega$ . We therefore have  $\mathbf{T}(r_\Omega, \theta, \phi, \omega) = 0$ , which can be written as (see eqs 6 and 7)

$$[{}_s\mathcal{Y}_l(r_\Omega, \omega)]_2 = [{}_s\mathcal{Y}_l(r_\Omega, \omega)]_4 = [{}_l\mathcal{Y}_l(r_\Omega, \omega)]_2 = 0, \quad (16)$$

where  $[\cdot]_i$  is the  $i$ th component of a vector. In the following, we omit indices  $l$  and  $s$  if expressions are the same for spheroidal and toroidal problems. The  $l$  index is also omitted in most of the expressions. Before applying boundary conditions, eq. (5) has four independent solutions for the spheroidal case and two for the toroidal case. Only two solutions are regular at the centre of the Earth for the spheroidal case and one for the toroidal case. The free surface boundary condition can only be met for a discrete set of eigenfrequencies,  $\{\omega_n, n \in \mathbb{N}\}$  where  $n$  is the radial order. For the toroidal case, an eigenfrequency is found when the traction at the surface vanishes and, for the spheroidal case, when the determinant of the traction component of the two remaining solutions at the surface vanishes. Finding all the eigenfrequencies to build a normal mode catalogue of a given spherically symmetric earth model requires a numerical scheme. The difficult part is to find all the eigenfrequencies without losing any of them, but once this is done, the normal

mode catalogue can, for example, be used to compute synthetic seismograms (see Dahlen & Tromp 1998 for a review). At this stage, we already have all the necessary ingredients to address our subject and there is no need to describe in more details the classical normal mode method theory to solve the wave equation in spherically symmetric earth models.

## 2.2 Non-periodic homogenization for the wave equation in layered media: summary of previous results

In this section, we summarize the results obtained by Capdeville & Marigo (2007) and Capdeville *et al.* (2010a) about non-periodic homogenization in layered media. For a given layered medium, this asymptotic method allows to compute the corresponding effective medium and correctors for a given minimum wavelength.

We first define a small parameter

$$\varepsilon_0 = \frac{\lambda_0}{\lambda_{min}}, \quad (17)$$

where  $\lambda_0$  is a characteristic size for which inhomogeneities smaller than  $\lambda_0$  are considered as microscopic and for which only effective properties are relevant to the wave equation,  $\lambda_{min}$  is the minimum wavelength of the wavefield. The  $\lambda_0$  value is user-defined and allows to define a lowpass filter operator such that, for any function of  $g(r)$ , the function

$$g^{\varepsilon_0*}(r) \equiv \mathcal{F}^{\varepsilon_0}(g)(r) \quad (18)$$

does not contain any variation faster than  $\lambda_0$ . The lowpass filter  $\mathcal{F}^{\varepsilon_0}$  can be defined from the Fourier orthogonal basis (see Appendix B) as we have done in our previous works (Capdeville & Marigo 2007; Capdeville *et al.* 2010a,b; Guillot *et al.* 2010). Nevertheless, other orthogonal bases can be used if they are more adapted to strong and slow velocity changes with depth in the Earth, inducing a strong change in the minimum wavelength from the bottom to the top of the earth model. A well-chosen basis allows to obtain a roughly constant  $\varepsilon_0$  everywhere despite the fact that  $\lambda_{min}$  changes with depth. See Appendix B for more details. We then introduce the microscopic space variable

$$y = \frac{r}{\varepsilon_0}, \quad (19)$$

where  $y$  is defined only on a limited segment  $[0, \lambda_{min}]$  and all quantities depending on  $y$  are extended to  $\mathbb{R}$  by periodicity. The principle of the homogenization procedure is to expand the solution  $\mathcal{Y}$  as a power series of  $\varepsilon_0$

$$\mathcal{Y}(r) = \mathcal{Y}^{\varepsilon_0,0}(r, y) + \varepsilon_0 \mathcal{Y}^{\varepsilon_0,1}(r, y) + \varepsilon_0^2 \mathcal{Y}^{\varepsilon_0,2}(r, y) + \dots \quad (20)$$

When  $\varepsilon_0 \rightarrow 0$ , any change in  $y$  induces a very small change in  $r$ . This leads to the separation of scales:  $y$  and  $r$  are treated as independent variables. Each term of the series  $\mathcal{Y}^{\varepsilon_0,i}$  depends upon the two space variables  $r$  and  $y$  (see Capdeville & Marigo 2007 and Capdeville *et al.* 2010a for details). The  $\mathcal{Y}^{\varepsilon_0,i}$  coefficients of the series are solutions of a coupled series of equations (Capdeville & Marigo 2007) that must be solved one by one.

$$\frac{\partial \mathcal{Y}^{\varepsilon_0,i}}{\partial r} + \frac{\partial \mathcal{Y}^{\varepsilon_0,i+1}}{\partial y} = \mathbf{S}^{\varepsilon_0} \mathcal{Y}^{\varepsilon_0,i} \quad (21)$$

with similar boundary conditions like (16), and where, in the non-periodic case,  $\mathbf{S}^{\varepsilon_0}(r, y)$  is built as follows:

$$\mathbf{S}^{\varepsilon_0}(r, y) = \mathcal{F}^{\varepsilon_0}(\mathbf{A})(r) + (\mathbf{I} - \mathcal{F}^{\varepsilon_0})(\mathbf{A})(\varepsilon_0 y), \quad (22)$$

where  $\mathbf{A}$  is defined in eq. (5) and  $\mathbf{I}$  is the identity operator. The last equation can be understood as follows: any smooth variation of  $\mathbf{A}$ ,

and therefore, any smooth variation of the parameter  $p_i$ , is assigned to the slow ( $r$ ) variable and any fast variation of  $\mathbf{A}$  is assigned to the fast variable ( $y$ ).

Solving the series of eq. (21), it is first shown that

$$\mathcal{Y}^0 = \langle \mathcal{Y}^0 \rangle, \quad (23)$$

where the cell average,  $\langle g \rangle$  for any function  $g(r, y)$ , is defined as

$$\langle g \rangle (r) \equiv \frac{1}{\lambda_{min}} \int_0^{\lambda_{min}} g(r, y) dy. \quad (24)$$

This implies that  $\mathcal{Y}^0$  does not depend upon the fast variable  $y$ . This is an important and intuitive result showing that, to the leading order, the solution to the wave equation is only affected in an average way by inhomogeneities much smaller than the minimum wavelength. Then, it is shown that  $\mathcal{Y}^0$ , which is the effective solution, is solution of the effective equation

$$\frac{\partial \mathcal{Y}^0}{\partial r} (r) = \mathbf{A}^{\varepsilon_0^*} (r) \mathcal{Y}^0 (r), \quad (25)$$

where the effective operator  $\mathbf{A}^{\varepsilon_0^*}$  is simply a lowpass filtered version of  $\mathbf{A}$

$$\mathbf{A}^{\varepsilon_0^*} = \langle \mathbf{S}^{\varepsilon_0} \rangle = \mathcal{F}^{\varepsilon_0} (\mathbf{A}) (r). \quad (26)$$

Still solving the series of eq. (21), it is then obtained, to the first order

$$\mathcal{Y} (r) = \mathcal{Y}^{\varepsilon_0, 0} \left( \mathbf{I} + \varepsilon_0 \mathbf{X}^{\varepsilon_0} \left( r, \frac{r}{\varepsilon_0} \right) \right) + O(\varepsilon_0^2), \quad (27)$$

where  $\mathbf{X}^{\varepsilon_0}$  is the non-periodic corrector, solution of  $\frac{\partial \mathbf{X}^{\varepsilon_0}}{\partial y} (y) = \mathbf{S}^{\varepsilon_0} (r, y) - \langle \mathbf{S}^{\varepsilon_0} (r, y) \rangle$  with periodic boundary conditions.  $\mathbf{X}^{\varepsilon_0}$  can be seen as the site effect: it is a local correction to the receiver. Finally, let us mention that, to the order first and second order of the asymptotic expansion (20), the boundary conditions are affected by the first- and second-order correctors (see Capdeville & Marigo 2007, 2008, for details). The new boundary condition is of type Dirichlet–Neumann

$$\mathbf{T}(r_{\Omega}, \theta, \phi) = D(\mathbf{u}(r_{\Omega}, \theta, \phi)), \quad (28)$$

where  $\mathbf{T}$  is the normal traction to the Earth surface,  $\mathbf{u}(r_{\Omega}, \theta, \phi)$  the displacement at the Earth surface and where the operator  $D$  depends on mainly five parameters related to integrals of non-linear combination of the elastic properties in the near surface. With good choices on the way is built the effective parameters near the free surface, the number of independent parameters controlling  $D$  can be reduced from five to two (Capdeville & Marigo 2007, 2008).

The effective form (26) and the fact that the  $\mathbf{A}^{\varepsilon_0^*}$  matrices linearly depend on the  $p_i$  parameters (see eqs A1 and A2) imply that the effective parameters  $p_i^{\varepsilon_0^*}$  are

$$p_i^{\varepsilon_0^*} (r) = \mathcal{F}^{\varepsilon_0} (p_i) (r). \quad (29)$$

The effective density and elastic parameters  $\rho^{\varepsilon_0^*}$ ,  $A^{\varepsilon_0^*}$ ,  $C^{\varepsilon_0^*}$ ,  $F^{\varepsilon_0^*}$ ,  $L^{\varepsilon_0^*}$  and  $N^{\varepsilon_0^*}$  are then deduced from the following relations:

$$\rho^{\varepsilon_0^*} = \mathcal{F}^{\varepsilon_0} (\rho), \quad (30)$$

$$\frac{1}{C^{\varepsilon_0^*}} = \mathcal{F}^{\varepsilon_0} \left( \frac{1}{C} \right), \quad (31)$$

$$\frac{1}{L^{\varepsilon_0^*}} = \mathcal{F}^{\varepsilon_0} \left( \frac{1}{L} \right), \quad (32)$$

$$A^{\varepsilon_0^*} - \frac{(F^{\varepsilon_0^*})^2}{C^{\varepsilon_0^*}} = \mathcal{F}^{\varepsilon_0} \left( A - \frac{F^2}{C} \right), \quad (33)$$

$$\frac{F^{\varepsilon_0^*}}{C^{\varepsilon_0^*}} = \mathcal{F}^{\varepsilon_0} \left( \frac{F}{C} \right), \quad (34)$$

$$N^{\varepsilon_0^*} = \mathcal{F}^{\varepsilon_0} (N), \quad (35)$$

which is the classical Backus (1962)'s result. A well-known consequence the above equations is that the effective medium is often anisotropic, even if the original medium is isotropic. Once the effective density and elastic parameters defined, one can easily deduce the effective velocities  $V_{ph}^{\varepsilon_0^*}$ ,  $V_{pv}^{\varepsilon_0^*}$ ,  $V_{sh}^{\varepsilon_0^*}$ ,  $V_{sv}^{\varepsilon_0^*}$  as well as  $\eta^{\varepsilon_0^*}$ .

The above development gives accurate results, the convergence is in  $\varepsilon_0^2$ , and for a small enough  $\varepsilon_0$ , the synthetic waveforms computed in the original medium and the effective one are the same up to the wanted precision. Refer to Capdeville & Marigo (2007) for test examples.

### 2.3 Non-periodic residual homogenization for the wave equation in layered media

In seismology, the inversion is often carried out with respect to a reference model. This reference model can contain small scales such as, for example, the PREM model (Dziewonski & Anderson 1981) which presents many elastic discontinuities. In the previous section, we have summarized an absolute homogenization for which no small scale is left in the effective medium. To account for the presence of a reference model, we describe here a modified homogenization, carried out with respect to a reference model, which we refer to as the residual homogenization. To do so, we assume that we have a reference earth model, defined by its density and elastic properties:  $(\rho_{ref}, A_{ref}, C_{ref}, F_{ref}, L_{ref}, N_{ref})$ , allowing to define the operators  $\mathbf{A}_{ref}$  based on (A1) and (A2). We do not make any particular assumption on the reference model and, for example, this model can be discontinuous. We now define the operator  $\mathbf{S}^{\varepsilon_0}$  (eq. 22) as

$$\mathbf{S}^{\varepsilon_0} (r, y) = \mathbf{A}_{ref} (r) + \mathcal{F}^{\varepsilon_0} (\mathbf{A} - \mathbf{A}_{ref}) (r) + (\mathbf{I} - \mathcal{F}^{\varepsilon_0}) (\mathbf{A} - \mathbf{A}_{ref}) (\varepsilon_0 y). \quad (36)$$

With this construction, all the results of Section 2.2 are still valid and we can find the effective operator, the receivers and sources correctors as well as the boundary conditions. For the effective properties, this new construction leads to

$$\mathbf{p}^{\varepsilon_0^*} (r) = \mathbf{p}_{ref} (r) + \mathcal{F}^{\varepsilon_0} (\mathbf{p} - \mathbf{p}_{ref}) (r), \quad (37)$$

and therefore

$$\rho^{\varepsilon_0^*} = \rho_{ref} + \mathcal{F}^{\varepsilon_0} (\rho - \rho_{ref}), \quad (38)$$

$$\frac{1}{C^{\varepsilon_0^*}} = \frac{1}{C_{ref}} + \mathcal{F}^{\varepsilon_0} \left( \frac{1}{C} - \frac{1}{C_{ref}} \right), \quad (39)$$

$$\frac{1}{L^{\varepsilon_0^*}} = \frac{1}{L_{ref}} + \mathcal{F}^{\varepsilon_0} \left( \frac{1}{L} - \frac{1}{L_{ref}} \right), \quad (40)$$

$$A^{\varepsilon_0^*} - \frac{(F^{\varepsilon_0^*})^2}{C^{\varepsilon_0^*}} = A_{ref} - \frac{F_{ref}^2}{C_{ref}} + \mathcal{F}^{\varepsilon_0} \left( A - A_{ref} - \frac{F^2}{C} + \frac{F_{ref}^2}{C_{ref}} \right), \quad (41)$$

$$\frac{F^{\varepsilon_0^*}}{C^{\varepsilon_0^*}} = \frac{F_{ref}}{C_{ref}} + \mathcal{F}^{\varepsilon_0} \left( \frac{F}{C} - \frac{F_{ref}}{C_{ref}} \right), \quad (42)$$

$$N^{\varepsilon_0*} = N_{\text{ref}} + \mathcal{F}^{\varepsilon_0} (N - N_{\text{ref}}) . \quad (43)$$

A consequence of such a construction is that the discontinuities and any other variations present in the reference model are still present in the effective medium. Only the residual between the real medium and the reference medium is homogenized and therefore smooth. In the following, we refer to

$$\delta \mathbf{p}(r) \equiv \mathbf{p}(r) - \mathbf{p}_{\text{ref}}(r) \quad (44)$$

as the Backus residual with respect to the reference model  $\mathbf{p}_{\text{ref}}$ . With such a definition, (37) can be rewritten

$$\mathbf{p}^{\varepsilon_0*}(r) = \mathbf{p}_{\text{ref}}(r) + \mathcal{F}^{\varepsilon_0} (\delta \mathbf{p})(r) . \quad (45)$$

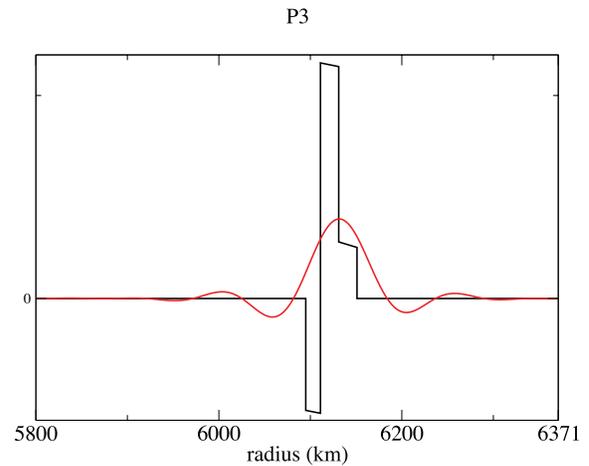
Note that (45) falls back to the classical result (29) if  $\mathbf{p}_{\text{ref}}(r)$  is chosen constant with  $r$ .

### 2.4 Residual homogenization examples

In this section, we present two validation examples of this procedure. More tests showing interesting effects of the residual homogenization are shown in Section 4. Our reference model will always be the PREM model. Two models are used:

- TEST1 model, with strong inhomogeneities relative to the reference model near the 220 km depth discontinuity (see Fig. 1, left-hand plot);
- TEST2 model, with the Moho discontinuity located at different depth from the reference model (see Fig. 1, right-hand plot).

Fig. 2 shows the  $\delta p_3(r)$  residual parameter, for the TEST1 model, as a function of the Earth radius (black line) and  $\mathcal{F}^{\varepsilon_0} (\delta p_3)(r)$  (red line) for  $\varepsilon_0 = 0.5$  and  $\lambda_{\text{min}} = 80$  km (with such parameters, the maximum frequency for synthetic seismograms with good accuracy would be of 1/40 Hz). It can be seen that for the residual parameter, the homogenization is trivial and involves only a linear lowpass filtering operation. Once the effective Backus residuals are obtained, using (39)–(43) and (8), the effective velocities and  $\eta^*$  parameters can easily be obtained. They are plotted in Fig. 3 for the TEST1 model and in Fig. 4 for the TEST2 model. More examples of the effects of the residual homogenization are shown in Section 4. Finally, Figs 5 and 6 show examples of traces computed in the residual effective model, for the TEST1 and TEST2 models, respectively, and compared with the reference solution (computed



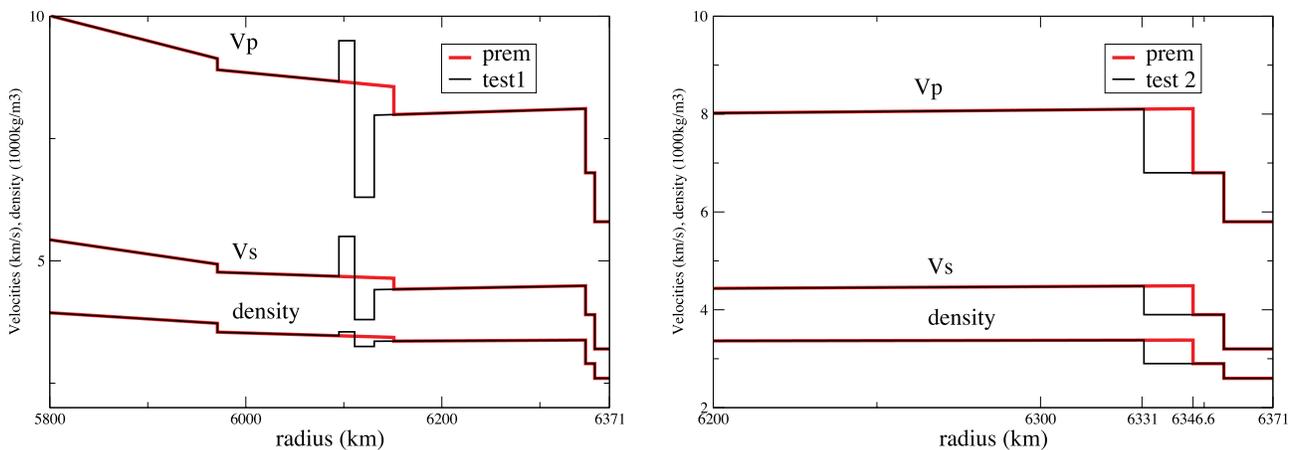
**Figure 2.**  $\delta p_3(r) = [\mathbf{p} - \mathbf{p}_{\text{ref}}]_3(r)$  for the model TEST1 (black line), PREM as a reference model and  $\mathcal{F}^{\varepsilon_0} (\delta p_3)(r)$  for the same model and  $\varepsilon_0 = 0.5$  (red line).

in the target models TEST1 and TEST2, not to be confused with the solutions obtained in reference model) and the solution obtained in the PREM model (the reference model). In both cases, the reference and homogenized traces show an excellent agreement.

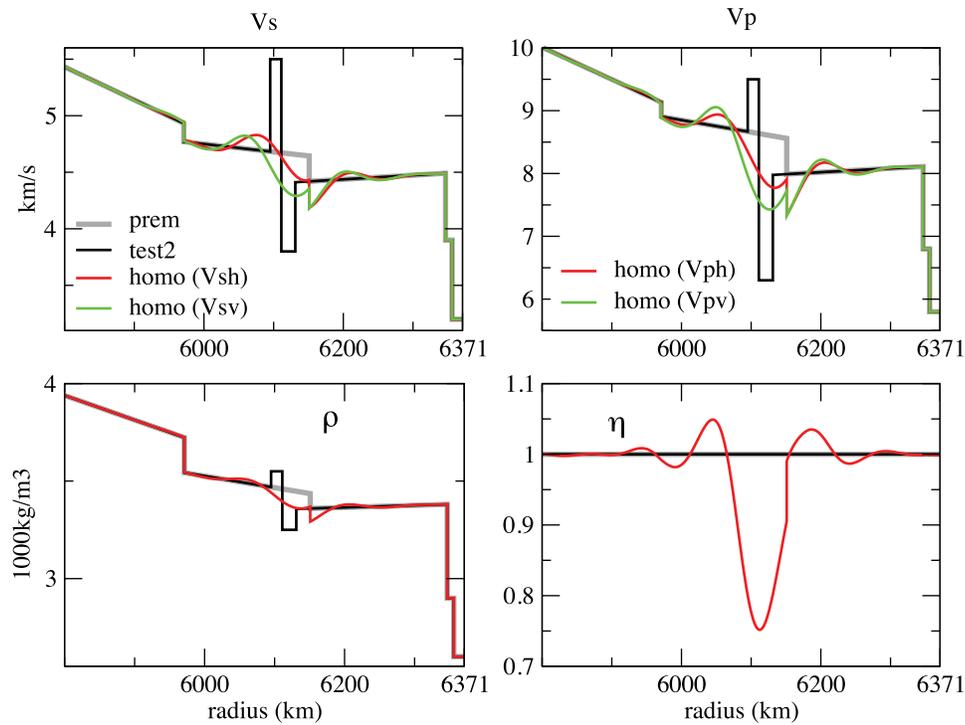
In this paper, we do not study the convergence with  $\varepsilon_0$ , but the result would be the same as the one shown in Capdeville & Marigo (2007, 2008), that is a convergence in  $\varepsilon_0^2$  of the effective solution towards the reference solution.

### 3 FULL-WAVEFORM INVERSION

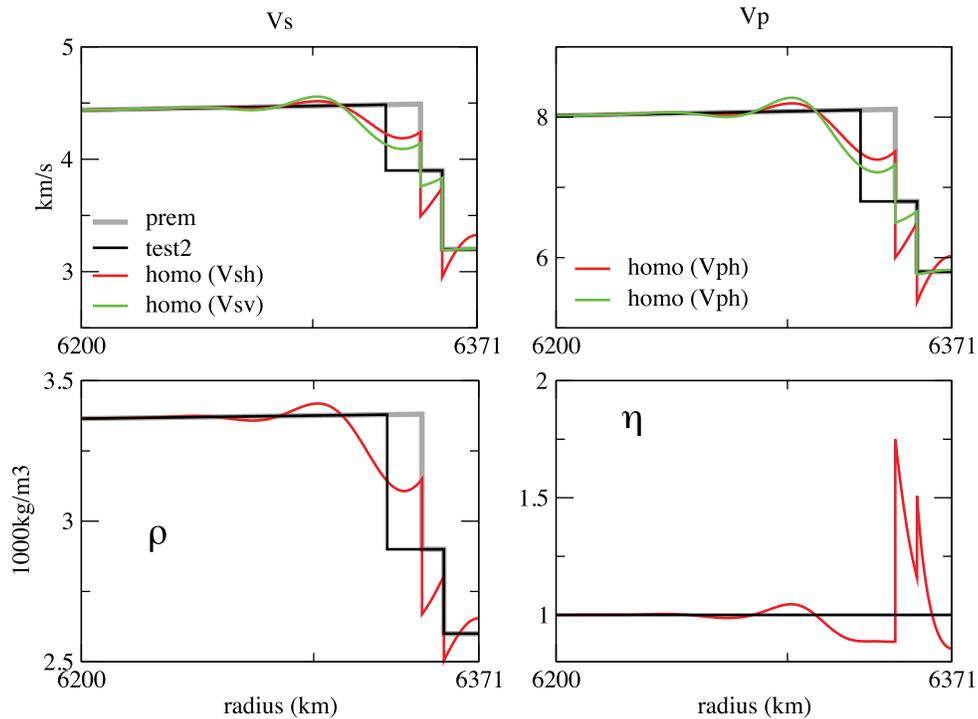
In this section, our objective is to numerically show that the effective medium obtained by residual homogenization (or classical homogenization) and the one obtained from a seismic inverse problem are the same. For this purpose and to make sure to use as much as possible information from the seismic traces, we introduce a full-waveform inversion method. We could have used other inversion methods, such as phase velocity inversion. Nevertheless, it is difficult with such methods to invert for the five elastic parameters plus density, which would have made our demonstration less obvious. The full-waveform inversion scheme developed here is not



**Figure 1.** TEST1 (left-hand plot) and TEST2 (right-hand plot) velocities and densities plotted with solid black lines as a function of the radius for the last kilometres close to the free surface. In red are plotted the velocities and density of an isotropic version of the PREM model for comparison.



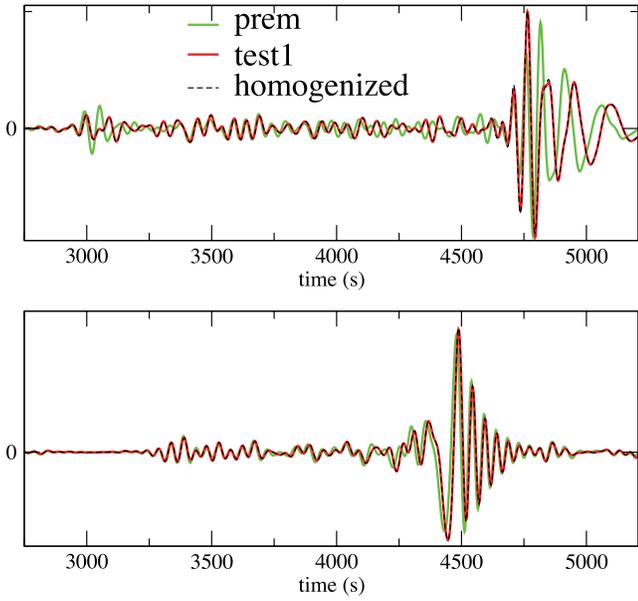
**Figure 3.**  $S$ -wave velocities (top left graph),  $P$ -wave velocities, (top right plot), density (bottom left plot) and  $\eta$  parameter (bottom right plot) as a function of  $r$  from PREM (grey line) and TEST1 models (black line). The horizontal (red lines) and vertical (green lines)  $S$ - and  $P$ -wave velocities, density (red line) and  $\eta$  parameter (red line) of the residual homogenized model of TEST1 with respect to PREM for  $\varepsilon_0 = 0.5$  are plotted.



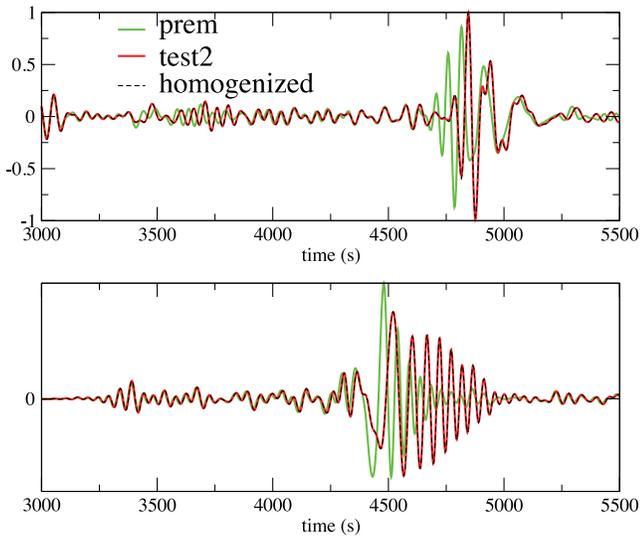
**Figure 4.**  $S$ -wave velocities (top left graph),  $P$ -wave velocities, (top right plot), density (bottom left plot) and  $\eta$  parameter (bottom right plot) as a function of  $r$  from PREM (grey line) and TEST2 models (black line). The horizontal (red lines) and vertical (green lines)  $S$ - and  $P$ -wave velocities, density (red line) and  $\eta$  parameter (red line) of the residual homogenized model of TEST2 with respect to PREM for  $\varepsilon_0 = 0.5$  are plotted.

intended to be applied to real data. Indeed, the assumption that the earth model is spherically symmetric will be heavily used and the Earth is not spherically symmetric. The spherically symmetric Earth approximation is often used for the phase velocity inversion, never-

theless amplitude is more sensitive to 3-D structure and our method would have little chance to be successful for signal amplitude due to a large theoretical error, at least when all the mode branches, from body waves to surface waves, are used. If only the fundamental



**Figure 5.** Vertical (top plot) and transverse (bottom plot) component seismograms for a 110 km depth source and an epicentral distance of  $130^\circ$  computed in PREM (green line), in TEST2 model (red line) and in the residual homogenized TEST1 model with respect to PREM for  $\varepsilon_0 = 0.5$  and with the order 2 boundary condition (dashed black line). The source has a maximum frequency of  $1/40$  Hz.



**Figure 6.** Vertical (top plot) and transverse (bottom plot) component seismograms for a 110 km depth source and an epicentral distance of  $130^\circ$  computed in PREM (green line), in TEST2 model (red line) and in the residual homogenized TEST2 model with respect to PREM for  $\varepsilon_0 = 0.25$  and with the order 2 boundary condition (dashed black line). The source has a maximum frequency of  $1/40$  Hz.

mode and a few harmonics are used, a real data application of the present method could be attempted.

Let  $\mathbf{g}$  be the forward modelling function that allows us to model the waveforms data ( $\mathbf{d}$ ) for a given set of model parameters ( $\mathbf{m}$ ):  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ . The inverse problem has to minimize the classical cost function  $\Phi$

$$\Phi(\mathbf{m}) = {}^T[\mathbf{g}(\mathbf{m}) - \mathbf{d}]\mathbf{C}_d^{-1}[\mathbf{g}(\mathbf{m}) - \mathbf{d}] + {}^T(\mathbf{m} - \mathbf{m}^0)\mathbf{C}_p^{-1}(\mathbf{m} - \mathbf{m}^0), \quad (46)$$

where  $\mathbf{m}^0$  is the *a priori* value of the model parameters (the starting model),  $\mathbf{C}_d$  and  $\mathbf{C}_p$  are the covariance matrices of data and model parameters, respectively. If  $\mathbf{g}$  is a non-linear function, the minimum, or the closest local minimum to the starting model, of the cost function  $\Phi$  can be found using the Gauss–Newton method iterative process (Tarantola & Valette 1982). Given  $\mathbf{m}^i$  the inverted model at iteration  $i$ , we can obtain model at the iteration  $i + 1$  as

$$\mathbf{m}^{i+1} = \mathbf{m}^i + ({}^T\mathbf{G}^i\mathbf{C}_d^{-1}\mathbf{G}^i + \mathbf{C}_p^{-1})^{-1} \times [{}^T\mathbf{G}^i\mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{g}(\mathbf{m}^i)) - \mathbf{C}_p^{-1}(\mathbf{m}^i - \mathbf{m}^0)], \quad (47)$$

where  $\mathbf{G}^i$  is the partial derivative matrix

$$\mathbf{G}^i = \left[ \frac{\partial \mathbf{g}(\mathbf{m})}{\partial \mathbf{m}} \right]_{\mathbf{m}=\mathbf{m}^i}. \quad (48)$$

In our case, the earth models are spherically symmetric, which allows to use efficiently the classical normal mode summation method to build  $\mathbf{g}$  with no specific approximation except the summation truncation, as well as the partial derivative matrix at each iteration  $\mathbf{G}^i$  with normal mode coupling, still with no specific approximation (see Appendix C).

To perform our exercise, we make some particular choices. First, if the covariance matrices  $\mathbf{C}_d$  and  $\mathbf{C}_p$  have an important meaning for linear inversions, this is less obvious for significantly non-linear cases such as the one considered here. We set  $\mathbf{C}_p^{-1} = \mathbf{0}$ , which implies that the final model  $\mathbf{m}^\infty$  can be as far as necessary from the starting model  $\mathbf{m}^0$ . Nevertheless, to avoid issues with the small eigenvalues of  ${}^T\mathbf{G}^i\mathbf{C}_d^{-1}\mathbf{G}^i$ , we add to it a damping diagonal matrix  $\lambda^i$ .  $\lambda^i$  is allowed to vary during the inversion and the damping is smaller and smaller with the increasing number of iterations. Furthermore, as it is often done for many seismic inversion methods (e.g. Bunks *et al.* 1995; Pratt 1999; Virieux & Operto 2009), to avoid to fall too quickly in a local minimum, the frequency band of the data changes with the iterations, starting with the low frequencies only and then increasing the frequency band little by little. This practice can be seen as a trick to avoid local minimum but, in our case, this operation has a special meaning that will be discussed in Section 5. To change the frequency band depending on the iteration number, we introduce a band limited data set

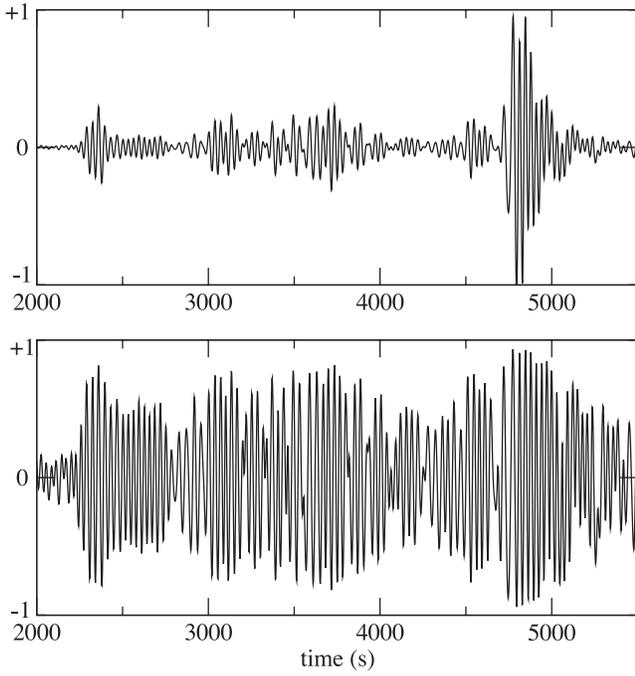
$$\mathbf{d}^i \equiv \mathbf{d} * w^i, \quad (49)$$

where  $w^i$  is a lowpass filter selecting the wanted frequency band for iteration  $i$  and  $*$  the temporal convolution. As the frequency band depends upon the iteration number, the function  $\mathbf{g}$  also depends upon  $i$  ( $\mathbf{g}^i(\mathbf{m}^i) \equiv \mathbf{g}(\mathbf{m}^i) * w^i$ ). To avoid the classical problem of relative amplitude between surface waves and body waves, the data are normalized by the inverse of its envelope plus a constant quantity. The data used at iteration number  $i$  can be written as

$$\tilde{\mathbf{d}}^i = \frac{\mathbf{d}^i}{e(\mathbf{d}^i) + c}, \quad (50)$$

where  $e(\cdot)$  denotes the envelope and  $c$  is a small value designed to avoid divisions by zero. This operation balances the amplitude and gives approximately the same weight to all seismic phases of the seismogram, as it can be seen on the example given in Fig. 7. From a notation point of view, it is convenient to hide this data normalization in the matrix  $\mathbf{C}_d$ , where, the envelope depending on the lowpass filter  $w^i$ , now also depends on  $i$ . This can be done as follows:

$$\mathbf{C}_d^i = {}^T(e(\mathbf{d}^i) + c)^{-1}\mathbf{C}_d(e(\mathbf{d}^i) + c)^{-1}. \quad (51)$$



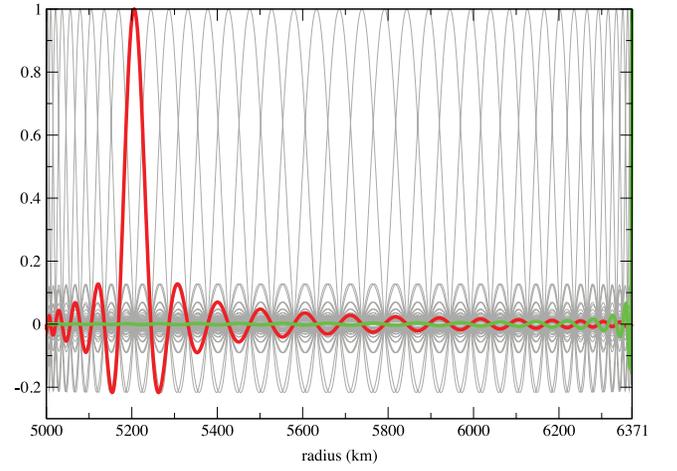
**Figure 7.** Example of the effect of the envelope normalization (50) on a vertical-component seismogram. In the top graph, is plotted the original seismogram and on the bottom graph is plotted the same seismogram after normalization.

Note that this operation is of minor importance for a synthetic test, even though it helps the convergence when artificial noise is added to the data. Let us mention that this normalization does not imply that the inversion scheme inverts for signal instantaneous phase instead of the full waveform. Indeed, the same normalization by the data envelope is also applied to the forward modelling operator  $\mathbf{g}^i$  and to the partial derivative matrix  $\mathbf{G}^i$  which implies that no partial derivative of the synthetic envelope are used. Again, this normalization plays a role only if noise or forward modelling errors are involved.

According to the above choices, eq. (47) can now be rewritten as

$$\mathbf{m}^{i+1} = \mathbf{m}^i + \left( {}^T\mathbf{G}^i \mathbf{C}_d^{i-1} \mathbf{G}^i + \lambda^i \right)^{-1} \left[ {}^T\mathbf{G}^i \mathbf{C}_d^{i-1} (\mathbf{d}^i - \mathbf{g}^i(\mathbf{m}^i)) \right]. \quad (52)$$

We now need to define the set of parameters  $\mathbf{m}^i$ . Classically, the model parameters are either velocities or slowness parametrized with spline functions or discrete layers versus depth. When layers are involved a vertical correlation length is introduced to reduce the number of independent parameters. The upper-mantle structure is dominantly recovered from the surface wave data which enable to resolve only  $S$ -wave velocity. When Rayleigh and Love wave data are simultaneously inverted,  $S$ -wave velocity radial anisotropy, corresponding to a VTI medium, can also be recovered. Dziewonski & Anderson (1981) describe this 1-D earth model with the  $A$ ,  $C$ ,  $F$ ,  $L$ , and  $N$  elastic parameters plus the density as a function of the radius. Because the model is overparametrized, correlations between parameters is further introduced (e.g. Montagner & Anderson 1989; Sebai *et al.* 2006; Ekström 2011; Debayle & Ricard 2012). Here, we invert for  $\delta\mathbf{p}(r)$ , the Backus residual vector with respect to a reference model (see eq. 44). For the radial discretization, we use the Lagrange polynomial interpolation associated with the Gauss–



**Figure 8.** Example of Lagrange polynomial basis associated with the Gauss–Lobatto–Legendre points for degree 40. Each polynomial is plotted in grey line, except the one associated with  $r_{10}$  ( $h_{10}^{40}(r)$ , red line) and the one associated with  $r_{40}$  ( $h_{40}^{40}(r)$ , green line).

Lobatto–Legendre (GLL) points spread in the inversion depth range (see Fig. 8 for an example). For a polynomial degree  $N$ , the full description of the model is given by the value of each parameters at the  $N + 1$  GLL depths.

$$\mathbf{m}^i = \left( \delta p_1^i(r_0), \delta p_2^i(r_0), \dots, \delta p_6^i(r_0), \delta p_1^i(r_1), \dots, \delta p_j^i(r_k), \dots, \delta p_6^i(r_N) \right), \quad (53)$$

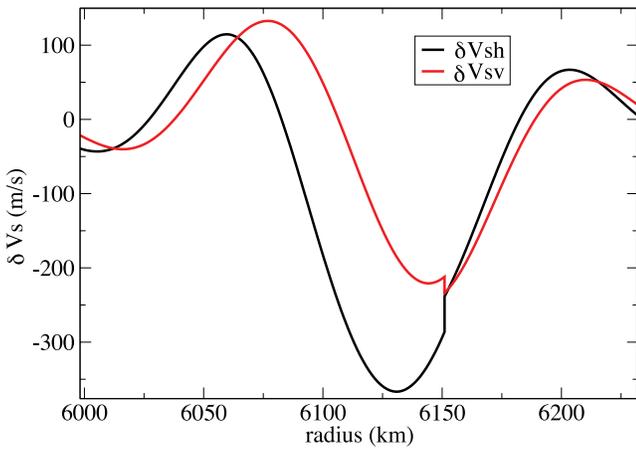
where  $\delta p_j^i(r_k)$  is the residual Backus parameter number  $j$  at iteration  $i$  and for the GLL radius  $r_k$ . If

$$h_k^N(r) \equiv \prod_{\substack{i=0, N \\ i \neq k}} \frac{r - r_i}{r_k - r_i} \quad (54)$$

is the Lagrange polynomial associated with the GLL point  $r_k$ , then the interpolation formula for the Backus residual parameters is

$$\delta\mathbf{p}^i(r) = \sum_{k=0}^N \delta\mathbf{p}^i(r_k) h_k^N(r). \quad (55)$$

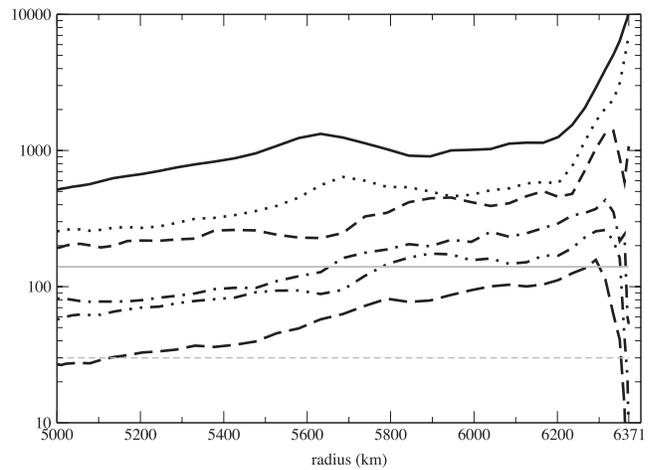
The Lagrange polynomial expansion used to radially expand the parameters  $\delta\mathbf{p}^i(r)$  forces the continuity and the smoothness of these parameters in the depth range of the inversion. Given that the reference model can contain discontinuities or scale smaller than the ones than can accurately be expanded on the Lagrange polynomial basis, constraining the residuals to be smooth is a strong assumption, which can be wrong for most parametrizations. Nevertheless, for the Backus residual parameters, this assumption is valid shown by eq. (45) (see Fig. 2 for an example). For other parametrizations than the Backus residuals, for example velocity residuals, this smoothness assumption is not valid. Indeed, the relation between the Backus residuals and most of the commonly used quantities, such as velocities, slowness, etc. are non-linear. As a consequence, if the reference model is discontinuous (or with any kind of fast variations), any other residual parameters than the Backus residuals are also discontinuous (see Fig. 9 for the  $V_s$  residuals for the TEST1 model example). There is therefore a clear advantage to use the Backus residuals as parameters for the inversion: it allows to correctly assume the smoothness of the inversion parameters with



**Figure 9.**  $V_{sh}$  and  $V_{sv}$  residuals with respect to PREM from the TEST1 homogenized model (see Fig. 3). These residuals are discontinuous at the PREM 220 km depth discontinuity.

depth and through the discontinuities of the reference model. Finally, let us mention that to be fully complete, we have to invert for the boundary term parameters mentioned eq. (28). In practice, we indeed have the possibility to invert for the boundary term parameters, but as shown by Capdeville & Marigo (2008), there is a strong trade-off between the boundary parameters and the elastic model values at the free surface. For the comparison between inverted and homogenized models we intend to make here, this trade-off can be a problem. To limit this problem, we make the choice to force the boundary parameters to be small and we have made sure that, for all the presented examples, the effect of the boundary term is small. Nevertheless, to study very shallow inhomogeneities, one should keep in mind to account for this trade-off.

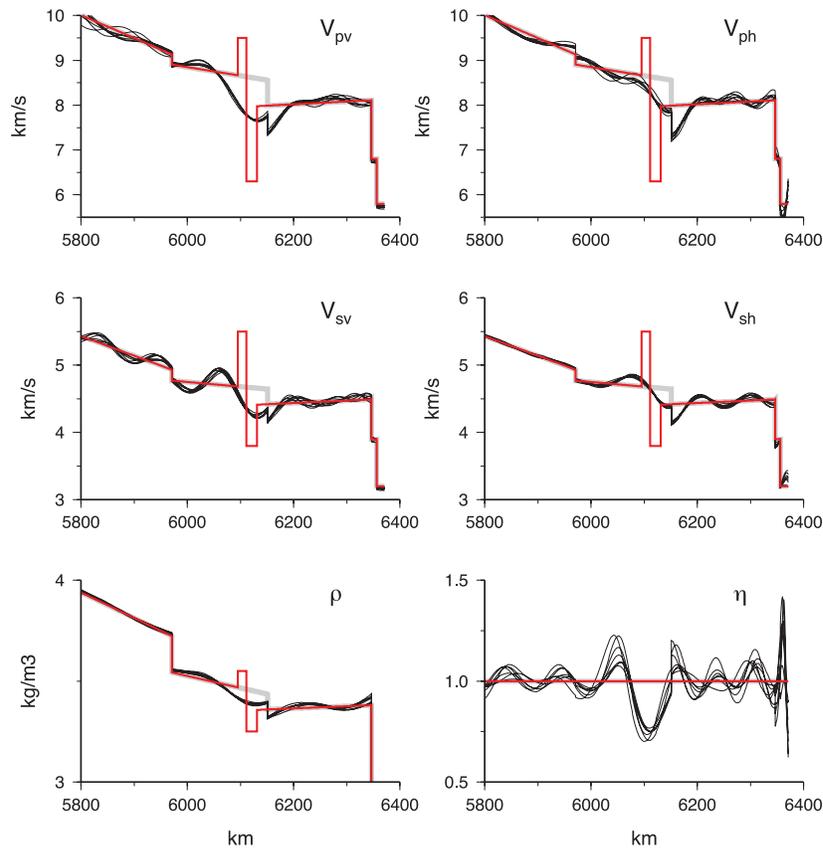
Before starting the inversion, to have an idea on how many parameters, among the six Backus residuals, can be inverted as a function of depth, we analyse the eigenvalues of the  $\mathbf{G}^0$  matrix computed in the reference model (here PREM) as a function of depth for a Lagrange polynomial degree  $N = 40$ , in the [5000–6371 km] depth range and for a minimum period of 33 s. To do so, we build a smaller  $\mathbf{G}^0$  matrix by keeping only the lines and columns corresponding to a single  $h_i^{40}(r)$ . This leads to a  $6 \times 6$  matrix for each GLL depth  $r_i$ , for which we can compute the eigenvalues and plot them as a function of  $r_i$  (see Fig. 10). This test does not give any information on the depth resolution, but it helps to identify potential trade-offs between the Backus residual parameters for a given depth. Because the problem is non-linear and because each depth should not be considered independently, this analysis can only be considered as indicative. In Fig. 10, we can guess how many parameters can be inverted per depth. Of course, the actual number of parameter that can indeed be inverted depends upon the noise level. In Fig. 10, two horizontal grey lines, corresponding to two examples of noise level, have been plotted. The dashed grey line corresponds to a low noise level and the solid grey line represents a higher noise level corresponding to a more realistic case. In Fig. 10, the actual vertical positions of the two grey lines is just indicative as a precise assessment of the noise level with respect to the eigenvalues is not possible for our non-linear global case. For the low noise level, it can be seen that below 5200 km and above 6350 km depth, we cannot retrieve the six parameters but in the range 5200–6350 km we have a good chance to retrieve the six parameters. For a real data inversion, the noise level may be much higher (corresponding, e.g. to the solid grey line in Fig. 10) and in such a case



**Figure 10.** Square root of the eigenvalues of submatrices of  $T\mathbf{G}^0\mathbf{C}_d^0{}^{-1}\mathbf{G}^0$  obtained by keeping only the lines and columns corresponding to a single  $h_i^N(r)$ , plotted as a function of  $r_i$  and normalized by the radius integral of  $h_i^N(r)$ . The reference model is PREM,  $N = 40$  and a minimum period of 33 s are used. Each black line corresponds to one of the six eigenvalue as function of depth. The two grey lines correspond to two noise level examples.

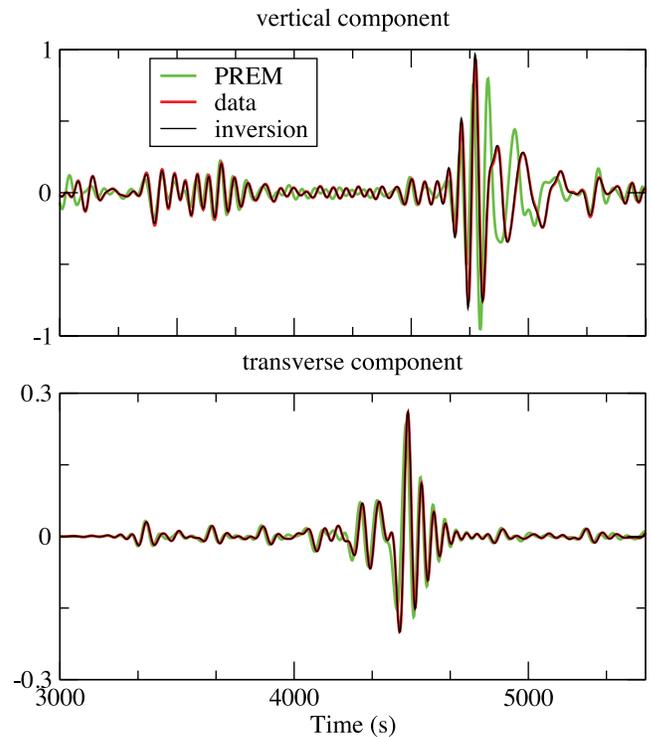
there is a little chance to retrieve the six parameters at any depth. For our synthetic tests, TEST1 model has been designed such that information to be retrieved is in the range [5800–6300 km]. Therefore, we can hope to access the six Backus residual parameters. For TEST2, the inhomogeneities are concentrated in the shallow [6300–6371 km] depth range and therefore we do not expect to retrieve all of the six Backus residual parameters, even for a low noise level.

To perform our synthetic tests, we first use the TEST1 model as a target model which means that the ‘data’ used for our inversion scheme described above are computed in the TEST1 earth model. We use a single three-component station and two earthquakes at the same surface location but for two different depths (10 and 700 km) corresponding to an epicentral distance of  $132^\circ$ . The data period band is [1000–33 s] and the Lagrange polynomial degree is  $N = 40$  in the [5000–6371 km] depth range. To avoid to commit an inversion crime, we add some random noise to the synthetic data to be inverted. The noise has the same frequency band as the data and is such that its standard deviation is 3 per cent of the standard deviation of the data. Such a noise is probably weak compared to a real case (it approximately corresponds to the low noise level horizontal dashed line shown in Fig. 10), but is large enough to avoid to use the forward modelling method error as an information and to commit an inversion crime. Moreover, local correctors at the source and receiver locations (see eq. 27) are not inverted for. This induces a few per cents theoretical error (mainly on amplitude) which also helps to prevent for an inversion crime. We intend to invert these full waveforms starting from the PREM model and with PREM as a reference model (the  $\mathbf{p}_{ref}$  model is therefore PREM and  $\mathbf{m}^0 = \mathbf{0}$ ). To have an idea of the robustness of the inversion results, we perform eight inversions with each time a different realization of the 3 per cent random noise added to the data. The raw results of the inversion are given in Fig. 11. The agreement between the inversion outputs (black lines in Fig. 11) and the TEST1 target model (red lines in Fig. 11) is not obvious. Nevertheless, computing a synthetic data in one of the inverted model for a receiver that has not been used for the inversion and comparing it with a reference synthetic seismogram computed in the TEST1 target model shows a very good

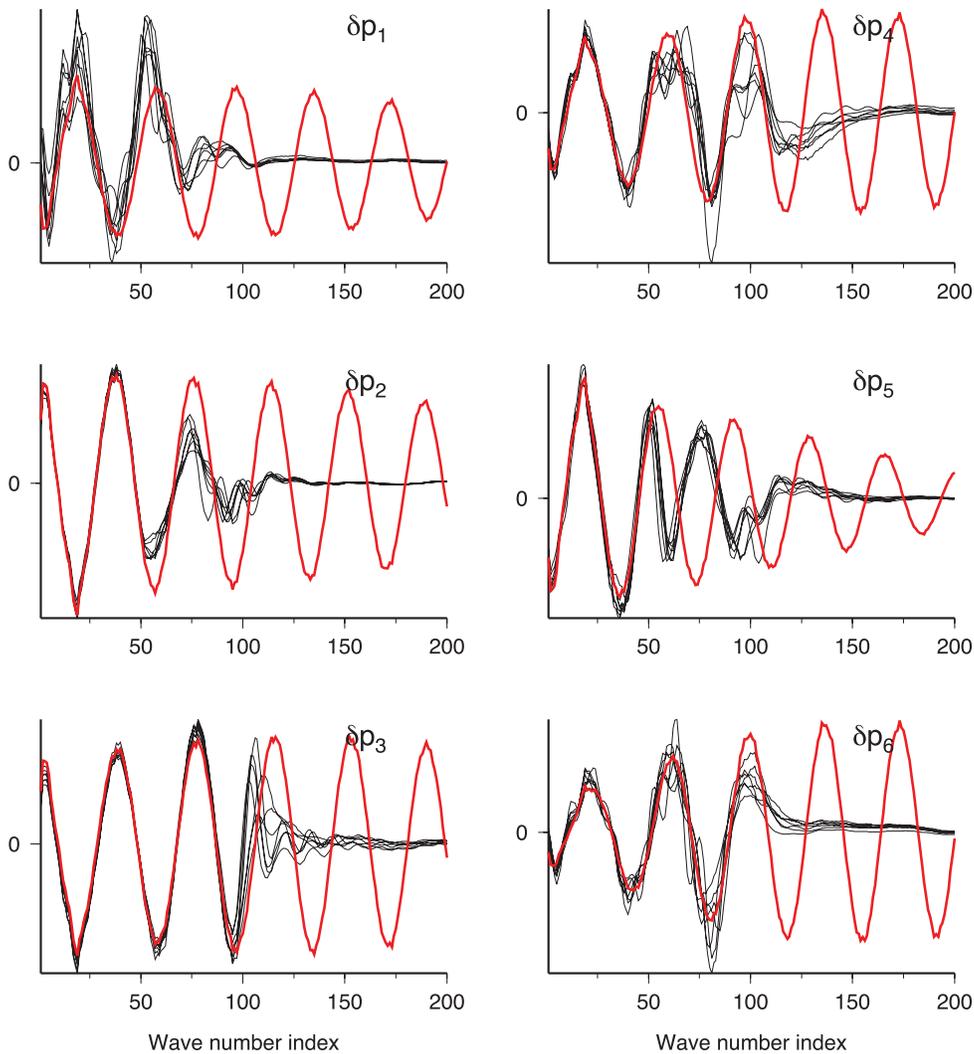


**Figure 11.** Raw results of the inversion. In grey, is plotted the reference model (PREM), in red the target model (TEST1) and in black the eight inversion results for eight different realizations of the 3 per cent random noise added to the data.

fit, as it can be seen in Fig. 12. So, even if the Fig. 11 comparison is not impressive, the inverted models are definitely not non-sense and are able to model accurately data in the [1000–33 s] frequency band, even for data that have not been used for the inversion. To perform a better comparison, we use the fact that, for spherically symmetric media, the important quantities are the Backus residual parameters in the spectral domain (see Appendix B for the radial spectral domain definition). Indeed, we have seen that the effective Backus residual parameters are obtained by lowpass filtering the residual parameters for the direct upscaling. The lowpass filtering being a simple product in the spectral domain (see Appendix B), a better comparison between the target and inverted models can be done in the spectral domain. The spectral domain comparison is done in Fig. 13. This time the agreement between the inversion results (black lines in Fig. 13) and the target model (red line in Fig. 13) is this time more obvious. As expected, the inversions give the correct answer up to an error bar that could be computed, but only for low wavenumbers. Indeed, below a wavenumber that depends upon the considered residual Backus parameter  $\delta p_i$ , the inverted and target models are in very good agreement, but above this number, they quickly diverge. On the conservative side, in Fig. 13, all parameters roughly agree for wavenumbers  $n \lesssim 70$ , which corresponds to a period higher than 17 s, that is about half of the lowest period of the inversion band. This means that the inversion is effective up to a resolution of roughly  $\lambda_{\min}/2$  where  $\lambda_{\min}$  is the minimum wavelength associated with the minimum period of the data used for the inversion. This is corresponding to the commonly admitted resolution power for most inversion methods. It can be seen that some parameters are very well recovered, but some, like the density ( $\delta p_1$ )



**Figure 12.** Seismograms computed in the output model of the inversion for a station that has not been used for the inversion (red line) to be compared with the data computed in the target model (black line) and in the starting model (green line). The epicentral distance is  $137^\circ$ .



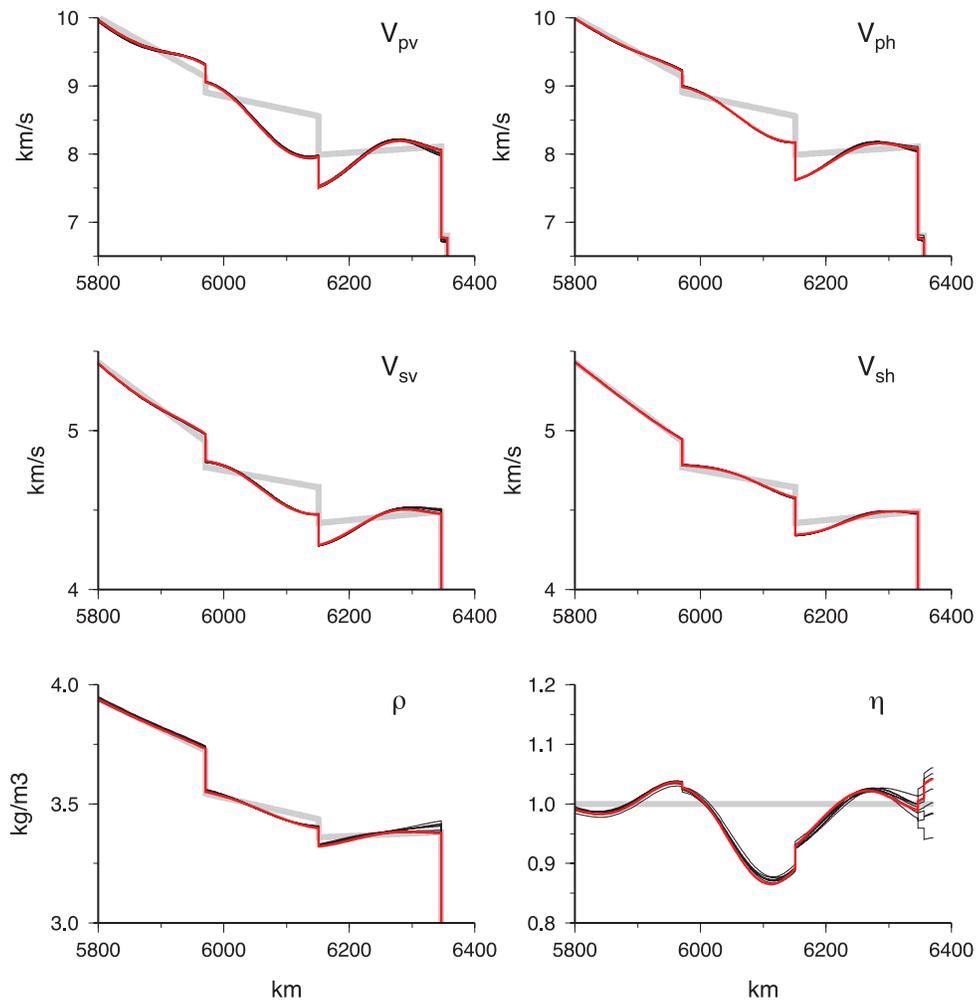
**Figure 13.** Raw results of the inversion for the TEST1 model in the spectral domain. In red is plotted the target model (TEST1) and in black the eight inversion results for eight different realizations of the 3 per cent random noise. For each plot, the PREM reference model would be the horizontal 0 line (not represented). To give a more physical idea of the horizontal axis values, modes  $n = 25, 50, 100$  and  $200$  have a period of, respectively, 47, 24, 12 and 6 s, which, for a wave speed of  $4 \text{ km s}^{-1}$ , would respectively correspond to a wavelength of 186, 95, 48 and 24 km.  $\delta p_1$  parameter is related to  $\rho$ ,  $\delta p_2$  to  $1/C$ ,  $\delta p_3$  to  $1/L$ ,  $\delta p_4$  to  $A - F^2/C$ ,  $\delta p_5$  to  $F/C$  and  $\delta p_6$  to  $N$ . See eqs (10)–(15) and (44) for the definition of the  $\delta p_i$  residuals.

are retrieved with less accuracy: the results of different inversions are scattered. Nevertheless, the different results gather around the target model spectrum which indicates that, despite a non-negligible error bar, the inversion is able to find all the six Backus residual parameters, including the correct density. To compare the models in the space domain, we conclude that the Backus residual spectral expansion of the inverted models need to be muted to zero for wavenumber beyond the resolution power of the inversion, and we therefore need to compare the inverted results with the target model homogenized to the same resolution. To do so, we homogenize the target and inverted models up to  $n = 70$  and the obtained results are plotted in Fig. 14. The match between the target model and the inverted models is this time very good, and we can conclude that they are the same up to a small error, which numerically shows that the inverted model is indeed the homogenized model.

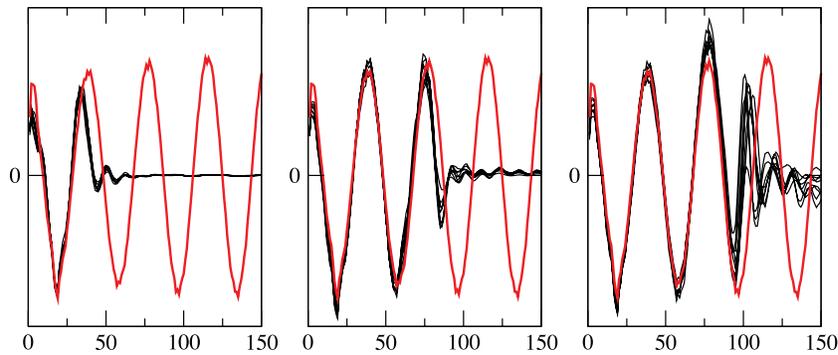
An interesting observation can be made plotting the inverted parameter  $\delta p_3$  in the spectral domain for different iterations of the inversion (see Fig. 15). As mentioned previously (see eq. 49), the frequency content of the data increase with the increasing iteration

number, and consequently, the resolution of the inversion results is also expected to increase with the increasing iteration number. It is indeed what is seen in Fig. 15. However, the most interesting observation that can be made from Fig. 15 is that, once inverted, the obtained low wavenumbers almost do not change while adding new frequencies. With a more sophisticated spatial parametrization, it would therefore be possible to invert only a new wavenumber range for each new frequency band added to the inversion while keeping unchanged the lower wavenumber obtained from previous iterations. Such an observation is only possible for the Backus residual parametrization, and, with such a parametrization the inversion is multiscale in the sense that adding frequencies to that data does not modify low wavenumber of the inverted models.

We conduct the same experiment for the TEST2 model. From the eigenvalue analysis shown in Fig. 10, three of the six eigenvalues are going to zero at the free surface. Knowing that the inhomogeneities of TEST2 model are focused in the crust, that is, close to the free surface for the considered wavelengths, we expect that it will not be possible to retrieve all of the six Backus residual



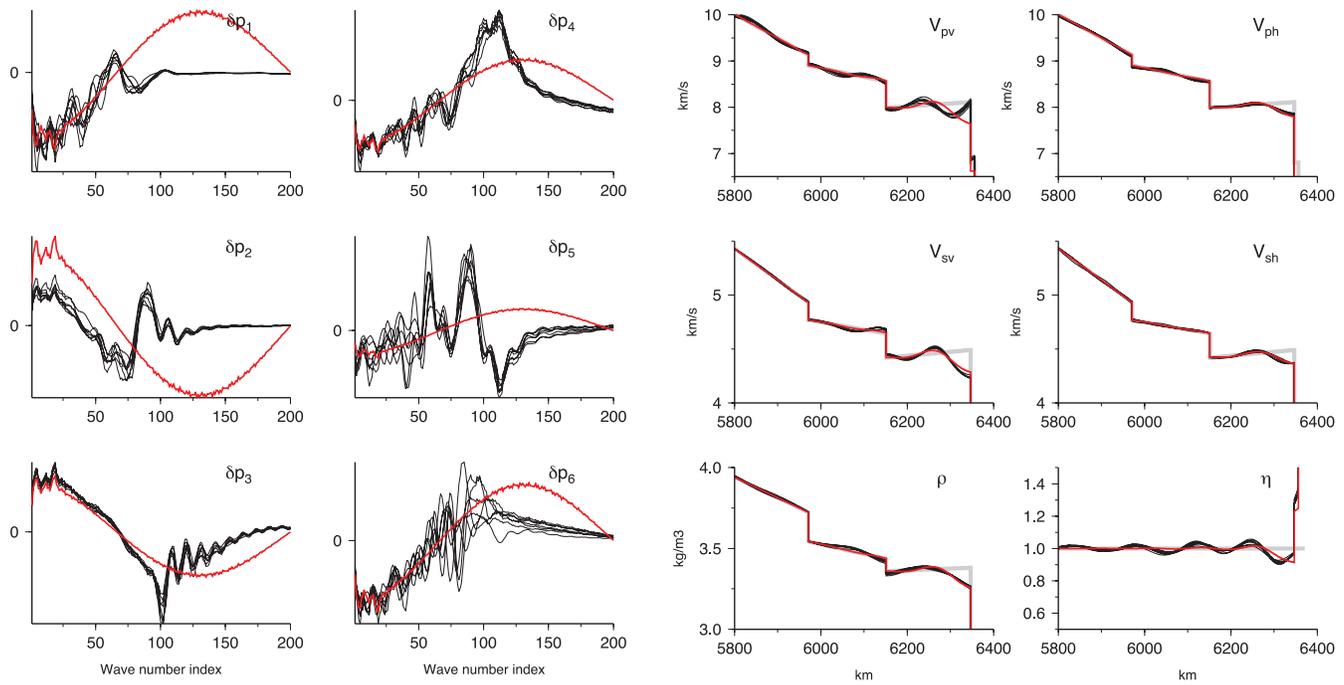
**Figure 14.** Upscaled results of the TEST1 inversion. In grey, is plotted the reference model (PREM), in red the target model (TEST1) and in black the eight inversion results for eight different realizations of the 3 per cent random noise.



**Figure 15.** Inversion results for the TEST1 model, for  $\delta p_3$  in the spectral domain for three different increasing iteration numbers of the inversion (and therefore three different increasing maximum frequencies of the data frequency band).

parameters. In Fig. 16 are plotted the results of the inversions, in the spectral domain (left-side six plots) and in the space domain, but homogenized down to  $n = 70$  (right-side six plots). As expected, the results accuracy is not as good as for the TEST1 case. Indeed, the different inversions are far more scattered for parameters  $\delta p_4$ ,  $\delta p_5$  and  $\delta p_6$  than for the TEST1 inversions. Moreover, the  $\delta p_2$  amplitude is way too low and we can consider that  $\delta p_2$  is not retrieved at all. Nevertheless, comparing the target and inverted models in the space domain as it has been done for TEST1

(Fig. 16, right-side six plots), the quality of the results is acceptable (the eight solutions gather around the true value), except for  $V_{pv}$  and  $\eta$ . The TEST2 numerical experience shows that, once again, the inversions retrieve the homogenized model, even if no specific treatment has been done to take into account the specific pattern of the shallow inhomogeneities. To improve a shallow inversion, one should take into account the strong trade-off between the shallow inhomogeneities (within roughly  $\lambda_{\min}/4$  distance to the free surface) and the homogenized Dirichlet–Neumann boundary condition



**Figure 16.** TEST2 inversion results. Left-side six plots: raw results of the inversion in the spectral domain (see eqs 10–15 and 44 for the definition of the  $\delta p_i$  residuals). In red is plotted the target model (TEST2) and in black the eight inversion results for eight different realizations of the 3 per cent random noise.  $\delta p_1$  parameter is related to  $\rho$ ,  $\delta p_2$  to  $1/C$ ,  $\delta p_3$  to  $1/L$ ,  $\delta p_4$  to  $A - F^2/C$ ,  $\delta p_5$  to  $F/C$  and  $\delta p_6$  to  $N$ . See eqs (10)–(15) and (44) for the definition of the  $\delta p_i$  residuals. Right-side six plots: upscaled results of the inversion. In grey, is plotted the reference model (PREM), in red the target model (TEST2) and in black the eight inversion results for eight different realizations of the 3 per cent random noise.

coefficients (Capdeville & Marigo 2008). A potential solution to the fact that only three parameters can be inverted for at the free surface, would be to invert for the two significant parameters of the Dirichlet–Neumann boundary condition and for the density while forcing the Backus residuals parameters at the free surface to zero. This would stabilize the inversion but would make the inversion interpretation difficult. We discuss further the interpretation issue in the conclusion section.

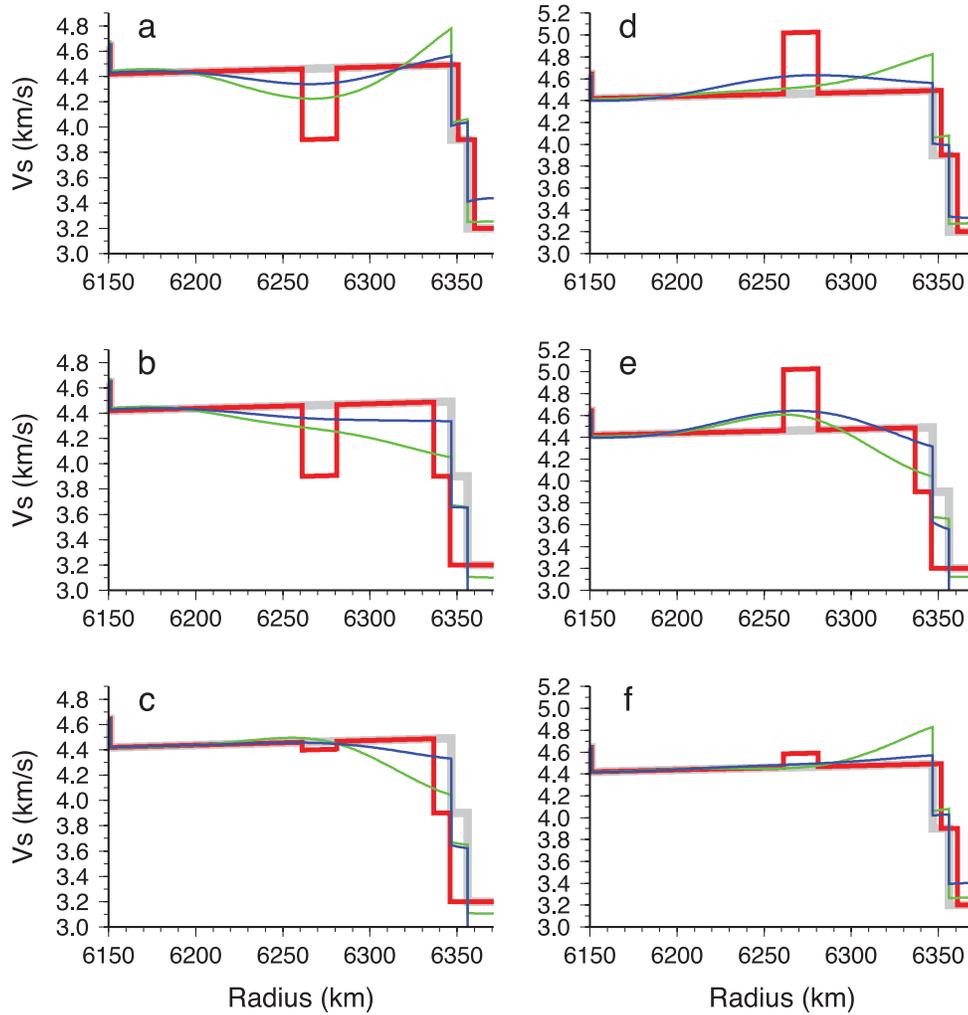
#### 4 RESIDUAL UPSCALING GALLERY

We have numerically shown in synthetic examples, that the effective medium and the inverted medium are at best the same, that is the inverted medium cannot be more accurate than the effective medium. We can therefore directly use the homogenization upscaling tool to infer what would be retrieved at best by an inversion without actually running the inversion. As an example, we show in Fig. 17 some potential effects of mislocated crust interface on the interpretation of inversion. The homogenized  $S$ -wave speed is plotted for six different target models (red line). Each target model has a low- or fast-velocity zone located around 100 km depth. The residual homogenization is performed with respect to PREM (grey line) in the same conditions as for the previous section (for a period cut-off of 33 s) and the resulting two homogenized wave speeds are shown (green:  $V_{sv}$ ; blue:  $V_{sh}$ ). We have chosen to show only  $V_s$  results to avoid any problems that could arise from the fact that a real inversion would have difficulties to recover all the parameters at the free surface (see TEST2 inversion test in the previous section). For each plotted case, the reference crust is mislocated, that is the target model crust is not the one of the reference model. For case (a) (Fig. 17 a), the crust interfaces mislocation does not change the apparent location of the 100 km depth low-velocity zone, but

creates an apparent high-velocity zone just below the Moho. For case (b), the localized low-velocity zone disappears and for case (c), we can even see a small-amplitude fast-velocity zone instead of a slow-velocity anomaly. Case (d) would be difficult to interpret: depending on which  $S$ -wave speed we are looking at, we can have a different interpretation. Finally, interpretation of cases (e) and (f) would be similar as the one of, respectively, cases (b) and (c). To conclude, the message of this test is that, even if an inversion gives a good result, the interpretation of this result can be non-trivial and misleading. Finally, let us mention that the similar issue of a wrong *a priori* crustal model have been studied by Bozdağ & Trampert (2007) and Ferreira *et al.* (2010) and the uses of apparent anisotropic crustal models to improve numerical efficiency has been studied by Fichtner & Igel (2008) and Lekić & Romanowicz (2011).

#### 5 DISCUSSION AND CONCLUSIONS

In this paper, we have first shown, in the non-periodic deterministic layered medium or spherically symmetric case, how to perform residual homogenization. This homogenization with respect to a reference model allows to perform more sophisticated operations than the classical homogenization. For example, it allows to homogenize only some interfaces of a discontinuous medium while keeping the other intact. Moreover, it allows to model what would produce an inversion relative to a non-homogeneous reference model. As usual, the residual homogenized model is anisotropic even if the fine scale model is isotropic. What is new and maybe puzzling is that the effective media now still have fine scale features like discontinuities and this is the case if the reference model contains fine scales. The tests performed for the forward problem show very a good accuracy. Note that, if we do not consider anelasticity for the homogenization



**Figure 17.** Grey: PREM (reference model), red: target model; green and blue: residual homogenized model (green:  $V_{sv}$ ; blue:  $V_{sh}$ ).

process in this paper, there is a no limitation to account for attenuation when solving the wave equation for the order 0 homogenized solution (the effective wave equation). If the attenuation model contains small scales, it can be upscaled the same way as for the elastic part. Nevertheless, we have not studied that aspect yet and a simple lowpass filtering of the attenuation model might be good enough for a start.

Then, using a full-waveform inversion method, we have numerically shown that an earth model resulting from an inversion is an homogenized model. Note that this result can be extended to any other inversion method based on frequency band limited data such as phase velocity inversion in the sense that such methods cannot recover more than an homogenized model (but it can recover less, because of an incomplete data coverage for example).

From the inversion perspective, this work raises some important issues. One point is that it helps to understand more clearly the necessity to increase little by little the frequency content of the data along the increasing iteration number of the non-linear least-square inversion algorithm, to avoid local minimum of the misfit functions. Indeed, the homogenization (or residual homogenization) principle shows that the elastic model ‘seen’ by the wavefield changes depending on the frequency cut-off applied to the seismic traces. The lower the frequency cut-off is, the smoother is the effective model. As an inversion can only recover what is ‘seen’ by the wavefield,

it implies that the target model depends upon the frequency. In other words, we can control the target model by tuning the cut-off frequency applied to the data. Knowing that for many cases, the reference models are good for very low frequencies, we can most of the time find a low enough frequency such that the target model is close enough to the reference model so that a linearized inversion scheme, such as the Gauss–Newton scheme, can easily converge without falling into a local minimum. So does it make the inversion result unique? So far not yet, mainly because of the free surface (there is a strong trade-off between the Dirichlet–Neumann boundary condition parameters mentioned in Section 2.2 and the shallow Backus residual parameters). However, it definitely reduces the non-uniqueness especially for deep inversion and for excellent data coverage.

An other result is that, from the technical point of view, we have shown that using a parametrization based on the Backus residual, the inversion is multiscale: inverting more and more higher frequency data enlarges more and more the wavenumber spectrum of the inverted model without changing the low wavenumbers obtained from smaller data frequency band. This is not the case using a classical parametrization, such as velocity residuals.

The next important point is the interpretation of the results. As we have seen, the inversion can only retrieve at best an homogenized model and this can be a problem for the interpretation of the results.

Indeed, the delocalization effect of the homogenization can lead to non-trivial and misleading effects that can make the interpretation difficult. Actually, the interpretation of the results in term of geological structures (discontinuities in our layered case) would require a separate inverse problem, allowing to include *a priori* informations. This inverse problem can be seen as a dehomogenization problem: for a given smooth tomographic model obtained with a long-period waveform inverse problem, what are the possible fine scale models fitting with some given *a priori* informations which, once homogenized, are the same as this smooth tomographic model? The problem has a high probability to be highly non-unique and one can guess that we will need to rely on Monte Carlo type of inversion methods (e.g. Bodin *et al.* 2012a,b). This dehomogenization inverse problem is important as it addresses the main target of seismic inversions: obtaining information of the underground properties.

This work and conclusions need to be extended to higher dimension problems (the now popular and challenging 2-D and 3-D full-waveform inversions developed by, e.g. Pratt *et al.* 1998; Capdeville *et al.* 2005; Tromp *et al.* 2005; Fichtner *et al.* 2009; Virieux & Operto 2009; Lekić & Romanowicz 2011, and many others) and it will be the purpose of future works.

## ACKNOWLEDGEMENTS

The authors would like to thank Andreas Fichtner for his positive review. This work was funded by the ANR blanche ‘mémé’ grant (ANR-10-BLAN-613 MEME) and the European ITN QUEST. Computations were performed on the CCIPL computer ‘Erdre’.

## REFERENCES

- Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Methods*, Freeman, San Francisco.
- Backus, G., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.
- Beucler, E., Stutzmann, E. & Montagner, J.-P., 2003. Surface-wave higher mode phase velocity measurements, using a roller coaster type algorithm, *J. geophys. Res.*, **108**(1), 289–307.
- Bodin, T., Capdeville, Y. & Romanowicz, B.A., 2012a. The inverse homogenization: incorporating discontinuities in smooth full waveform inversion models, Abstract S31C-02 presented at 2012 Fall Meeting AGU, 3–7 Dec, San Francisco, Calif.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012b. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.*, **117**, B02301, doi:10.1029/2011JB008560.
- Bozdag, E. & Trampert, J., 2007. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**(3), 1066–1082.
- Bunks, C., Saleck, F., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.
- Capdeville, Y., 2005. An efficient Born normal mode method to compute sensitivity kernels and synthetic seismograms in the Earth, *Geophys. J. Int.*, **163**, 639–654.
- Capdeville, Y. & Marigo, J.J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophys. J. Int.*, **170**, 823–838.
- Capdeville, Y. & Marigo, J.J., 2008. Shallow layer correction for spectral element like methods, *Geophys. J. Int.*, **172**, 1135–1150.
- Capdeville, Y., Stutzmann, E. & Montagner, J.P., 2000. Effect of a plume on long period surface waves computed with normal modes coupling, *Phys. Earth planet. Inter.*, **119**, 57–74.
- Capdeville, Y., Gung, Y. & Romanowicz, B., 2005. Towards global earth tomography using the spectral element method: a technique based on source stacking, *Geophys. J. Int.*, **162**, 541–554.
- Capdeville, Y., Guillot, L. & Marigo, J.J., 2010a. 1-D non periodic homogenization for the wave equation, *Geophys. J. Int.*, **181**, 897–910.
- Capdeville, Y., Guillot, L. & Marigo, J.J., 2010b. 2D nonperiodic homogenization to upscale elastic media for *P-SV* waves, *Geophys. J. Int.*, **182**, 903–922.
- Cara, M. & Lévêque, J., 1987. Waveform inversion using secondary observables, *Geophys. Res. Lett.*, **14**(10), 1046–1049.
- Cara, M., Nolet, G. & Necessian, A., 1980. New inferences from higher mode data in western Europe and northern Eurasia, *Geophys. J. R. astr. Soc.*, **61**(3), 459–478.
- Cupillard, P., 2008. Simulation par la méthode des éléments spectraux des formes d’onde obtenues par corrélation de bruit sismique, *PhD thesis*, Université Paris 7 - Denis Diderot, Institut de Physique du Globe de Paris.
- Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, NJ.
- Debayle, E. & Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher Rayleigh mode measurements, *J. geophys. Res.*, **117**, B10308, doi:10.1029/2012JB009288.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**, 297–356.
- Ekström, G., 2011. A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25–250 s, *Geophys. J. Int.*, **187**, 1668–1686.
- Ekström, G., Tromp, J. & Larson, E., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.*, **102**, 8137–8157.
- Ferreira, A., Woodhouse, J., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. geophys. Res.*, **115**(B4), B04313, doi:10.1029/2009JB006716.
- Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models—a technique based on non-linear dispersion curve matching (DCM), *Geophys. J. Int.*, **173**(2), 519–533.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.P., 2009. Full waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, **179**, 1703–1725.
- Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villasenor, A., 2013. Multi-scale full waveform inversion, *Geophys. J. Int.*, 10.1093/gji/ggt118.
- Guillot, L., Capdeville, Y. & Marigo, J.J., 2010. 2-D non periodic homogenization for the SH wave equation, *Geophys. J. Int.*, **182**, 1438–1454.
- Lebedev, S. & van der Hilst, R., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and *S*-wave forms, *Geophys. J. Int.*, **173**(2), 505–518.
- Lekić, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.*, **185**(2), 799–831.
- Li, X.B. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.*, **101**, 11 245–22 271.
- Li, X.D. & Tanimoto, T., 1993. Waveforms of long-period body waves in slightly aspherical Earth model, *Geophys. J. Int.*, **112**, 92–102.
- Lognonné, P. & Romanowicz, B., 1990. Modelling of coupled normal modes of the Earth: the spectral method, *Geophys. J. Int.*, **102**, 365–395.
- Montagner, J. & Anderson, D., 1989. Constrained reference mantle model, *Phys. Earth planet. Inter.*, **58**(2), 205–227.
- Montagner, J. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.*, **95**(B4), 4797–4819.
- Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. & Hung, S., 2004. Finite-frequency tomography reveals a variety of mantle plumes, *Science*, **303**, 338–343.
- Nakanishi, I. & Anderson, D., 1982. Worldwide distribution of group velocity of mantle Rayleigh waves as determined by spherical harmonic inversion, *Bull. seism. Soc. Am.*, **72**(4), 1185–1194.
- Nishida, K., Montagner, J. & Kawakatsu, H., 2009. Global surface wave tomography using seismic hum, *Science*, **326**(5949), 112–112.
- Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.*, **95**(B6), 8499–8512.

- Pratt, R., 1999. Seismic waveform inversion in the frequency domain, part 1: theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901.
- Pratt, R., Shin, C. & Hicks, G., 1998. Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.
- Romanowicz, B., Gung, M.P.Y. & Capdeville, Y., 2008. On the computation of long period seismograms in a 3-D Earth using normal mode based approximations, *Geophys. J. Int.*, **175**, 520–536.
- Schimmel, M., Stutzmann, E. & Gallart, J., 2011. Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale, *Geophys. J. Int.*, **184**(1), 494–506.
- Sebai, A., Stutzmann, E., Montagner, J., Sicilia, D. & Beucler, E., 2006. Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography, *Phys. Earth planet. Inter.*, **155**(1), 48–62.
- Shapiro, N., Campillo, M., Stehly, L. & Ritzwoller, M., 2005. High resolution surface wave tomography from ambient seismic noise, *Science*, **307**, 1615–1618.
- Stutzmann, E. & Montagner, J.-P., 1993. An inverse technique for retrieving higher mode phase velocity and mantle structure, *Geophys. J. Int.*, **113**(3), 669–683.
- Takeuchi, H. & Saito, M., 1972. Seismic surface waves, *Methods Comput. Phys.*, **11**, 217–295.
- Tanimoto, T., 1984. A simple derivation of the formula to calculate synthetic long period seismograms in heterogeneous earth by normal mode summation, *Geophys. J. R. astr. Soc.*, **77**, 275–278.
- Tanimoto, T., 1986. Free oscillations of a slightly anisotropic earth, *Geophys. J. R. astr. Soc.*, **87**, 493–517.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, **20**, 219–232.
- Trampert, J. & Woodhouse, J., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.*, **122**(2), 675–690.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
- van Heijst, H. & Woodhouse, J., 1997. Measuring surface-wave overtone phase velocities using a mode-branch stripping technique, *Geophys. J. Int.*, **131**, 209–230.
- Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics*, **75**(6), WCC127–WCC152.
- Woodhouse, J. & Dahlen, F., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. R. astr. Soc.*, **53**, 335–354.
- Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized earth model, *Geophys. J. R. astr. Soc.*, **78**, 641–660.
- Yoshizawa, K. & Kennett, B., 2002. Non-linear waveform inversion for surface waves with a neighbourhood algorithm application to multimode dispersion measurements, *Geophys. J. Int.*, **149**(1), 118–133.

## APPENDIX A: A MATRICES IN SPHERICALLY LAYERED EARTH

For the spheroidal case, we have

$${}_s\mathbf{A}_l^e(r, \omega) = \begin{pmatrix} d/r & 1/C & e_l/r & 0 \\ -\rho\omega^2 + a & -d/r & -a\gamma_l/2 & \gamma_l/r \\ -\gamma_l/r & 0 & 2/r & 1/L \\ a\gamma_l/2 & -e_l/r & -\rho\omega^2 + b_l & -2/r \end{pmatrix} \quad (\text{A1})$$

with

$$a = \frac{4}{r^2} \left( A - \frac{F^2}{C} - N \right), \quad b_l = \frac{\gamma_l^2}{r^2} \left( A - \frac{F^2}{C} \right) - \frac{2N}{r^2}, \\ d = 1 - \frac{2F}{C}, \quad e_l = \gamma_l \frac{F}{C}.$$

In the toroidal case, we have

$${}_t\mathbf{A}_l^e(r, \omega) = \begin{pmatrix} 2/r & 1/L \\ -\rho\omega^2 + \Omega_l N/r^2 & -2/r \end{pmatrix} \quad (\text{A2})$$

with  $\Omega_l = (l-1)(l+2)$ .

From these expressions, one can see that the averaged quantities required to compute  $\langle \mathbf{S} \rangle$  are the Backus  $p_i$  parameters, that is  $\langle \frac{1}{C} \rangle$ ,  $\langle \frac{1}{L} \rangle$ ,  $\langle A - \frac{F^2}{C} \rangle$ ,  $\langle \frac{F}{C} \rangle$  and  $\langle N \rangle$  which is different from the average of the  $A, C, F, L, N$  elastic parameters. This is the same result as that obtained by Backus (1962).

## APPENDIX B: SPATIAL FILTERING AND SPECTRAL DOMAIN

In this appendix, we define more precisely the lowpass filter operator  $\mathcal{F}^{\varepsilon_0}(\cdot)$  used along the paper. We define  $B = \{u_n(r), n \in \mathbb{Z}\}$  a set of orthogonal functions that can be considered as a basis for any properties of the Earth on the radius segment  $R = [r_a, r_\Omega]$ .  $r_a$  is a radius corresponding to a depth below which the waveform are not significantly sensitive to the elastic properties of the earth model. Let  $g(r)$  be a function of  $r$  on  $R$  [ $g$  can for example be any of the Backus residual parameters  $\delta p_i(r)$ ],  $B$  being an orthonormal basis for such a function, we can write

$$g(r) = \sum_n g_n u_n(r), \quad (\text{B1})$$

where

$$g_n = \int_R g(r) u_n^*(r) dr, \quad (\text{B2})$$

where  $*$  is the complex conjugate. The  $g_n$  coefficients constitute the radial spectral representation of  $g$ . Note that  $dr$  in (B2) might be replaced by  $rdr$  to account for the sphericity, but it is of little effect knowing that the considered depth range is small with respect to  $r_\Omega$ . For a given  $\lambda_{\min}$  and for some Basis  $B$  (typically Fourier basis), we can precisely link an  $\varepsilon_0$  to an  $n_0$  such that  $u_{n_0}$  oscillate with characteristic wavelength of value  $\lambda_0$ . In that case, we simply have

$$\mathcal{F}^{\varepsilon_0}(g)(r) = \sum_{n, |n| \leq n_0} g_n u_n(r). \quad (\text{B3})$$

In practice, such a brutal cut-off is not a good idea because it delocalizes to much the properties. To avoid this drawback, we use a smoother transition to zero in the spectral model introducing a cosine taper

$$w_{n_1, n_2}(n) = \begin{cases} 1 & \text{for } |n| \leq n_1, \\ \frac{1}{2} \left( 1 + \cos \left( \pi \frac{|n| - n_1}{n_2 - n_1} \right) \right) & \text{for } |n| \in ]n_1, n_2], \\ 0 & \text{for } |n| \geq n_2, \end{cases} \quad (\text{B4})$$

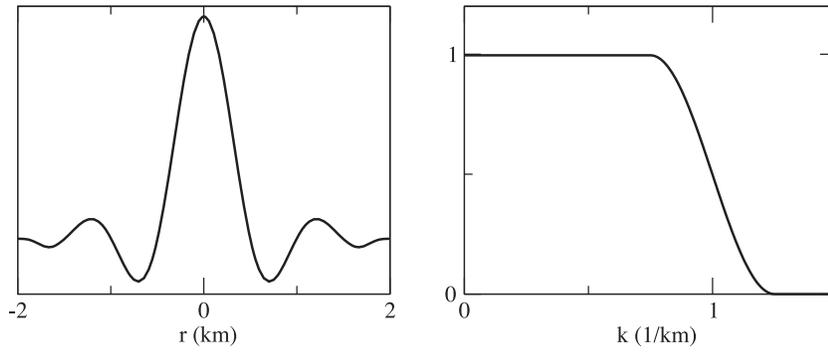
where  $n_1$  and  $n_2$  are two values close to  $n_0$ . We use

$$\mathcal{F}^{\varepsilon_0}(g)(r) = \sum_n g_n u_n(r) w_{n_1, n_2}(n). \quad (\text{B5})$$

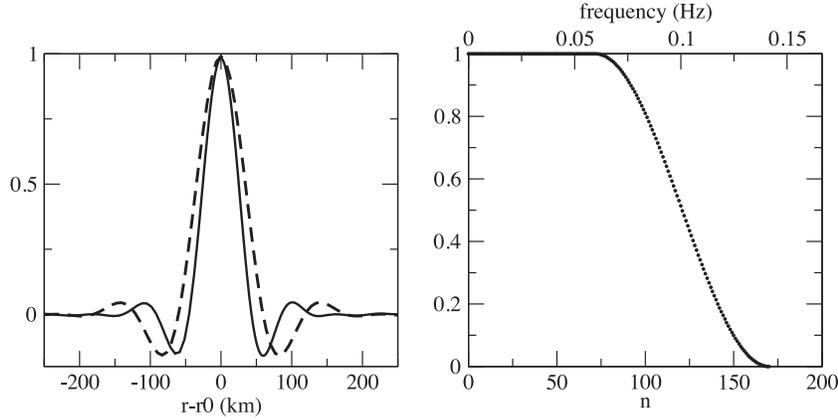
A natural basis  $B$  is the Fourier basis. In such case, we use

$$u_n(r) = \cos \left( n\pi \frac{r_\Omega - r}{r_\Omega - r_a} \right). \quad (\text{B6})$$

Note the fact that we use cosine Fourier basis instead of the classical complex exponential basis is due to the fact that we extend the  $g$  function for  $r > r_\Omega$  by the symmetric of  $g$  with respect to  $r_\Omega$ . For



**Figure B1.** Example Fourier wavelet  $w_{n_1, n_2}(r)$  (left-hand plot) and the corresponding  $w_{n_1, n_2}(k)$  where  $k = n/(r_\Omega - r_a)$  for a  $n_0/(r_\Omega - r_a) \simeq 1 \text{ km}^{-1}$ .



**Figure B2.** Left-hand plot: two examples of normal wavelet  $w_{n_1, n_2, r_0}$  (see eq. B10) for two values of  $r_0$ , 6000 km (solid line) and 5000 km (dashed line), for  $n_1 = 77$  and  $n_2 = 170$  and for a normal basis computed in a smooth version of PREM. It can be seen that for the same wavenumber cut-off, the wavelet is wider at depth. The wavelet is also slightly asymmetric. In the right plot is shown the filtering operator  $w_{n_1, n_2}(n)$  as a function of  $n$  (bottom axis) and as a function of the corresponding eigenfrequency value (top axis).

such symmetric functions, a cosine expansion is enough. For the Fourier case, we can compute the space domain wavelet

$$w_{n_1, n_2}(r) = \sum_n u_n(r) w_{n_1, n_2}(n), \quad (\text{B7})$$

and in such a case, the filtering operator can be written as

$$\mathcal{F}^{\varepsilon_0}(g)(r) = g * w_{n_1, n_2}(r), \quad (\text{B8})$$

where  $*$  is here the spatial convolution. An example of wavelet  $w_{n_1, n_2}$  is given in Fig. B1.

For many media, like all realistic spherically symmetric global earth models, the minimum wavelength strongly varies with depth. This is due to the increase of the average wave speed with depth. From the homogenization point of view, we can always choose the smallest minimum wavelength to measure  $\varepsilon_0$ , but this is not optimum. To account for the strong change of minimum wavelength with depth, we can choose a different set of function  $B$  than the Fourier basis. Remembering that the Fourier functions are solutions of the wave equation in a homogeneous medium, we can use solutions of the radial wave equations solved in a very smooth version of the actual earth model as the orthogonal function  $u_n$ . We therefore use

$$u_n(r) = r U_0(r, n \omega_0), \quad (\text{B9})$$

where the  $U_0$  is the radial eigenfunction defined in eq. (6),  $n \omega_0$  the infrequency of radial order  $n$  computed for  $l = 0$  and in a smooth earth model roughly reproducing the variation of the  $S$ -wave speed. The smooth model does not need to be precisely defined and any

smooth model roughly reproducing the wavelength changes with depth is acceptable. In such a case, the action of the filter  $\mathcal{F}^{\varepsilon_0}(\cdot)$  cannot be computed as a convolution like for the Fourier case, nevertheless, we can plot the space representation of the wavelet  $w_{n_1, n_2}$  for different depth. To do so, we compute

$$w_{n_1, n_2, r_0}(r) = \sum_n u_n(r) w_{n_1, n_2}(n) d_{r_0, n}, \quad (\text{B10})$$

where

$$d_{r_0, n} = \int_R \delta(r - r_0) u_n(r) dr. \quad (\text{B11})$$

Two examples of such wavelets are plotted in Fig. B2. We refer to this way of performing the lowpass filtering as the ‘normal mode filtering’.

## APPENDIX C: PARTIAL DERIVATIVE COMPUTATION WITH NORMAL MODES

For spherically symmetric earth models, the partial derivatives for the displacement solution to the wave equation with respect to the Earth’s elastic and density can be efficiently computed from the Earth normal mode method (this subject has been addressed by many authors, e.g. Woodhouse & Dahlen 1978; Tanimoto 1984; Lognonné & Romanowicz 1990; Capdeville *et al.* 2000; Romanowicz *et al.* 2008). This appendix is specifically based on Capdeville (2005) and we keep the same notations. Considering only a perturbation of the anelastic tensor  $\delta \mathbf{c}$ , the displacement perturbation can

be rewritten (see Capdeville 2005) as

$$\delta \mathbf{u}(\mathbf{r}_r, \omega) \cdot \mathbf{v} = \int_V \left[ \sum_K \frac{\boldsymbol{\epsilon}_K^*(\mathbf{r}) R_K}{(\omega_K^2 - \omega^2)} \right] : \delta \mathbf{c}(\mathbf{r}_d) : \left[ \sum_K \frac{\boldsymbol{\epsilon}_K(\mathbf{r}) S_K}{i\omega(\omega_K^2 - \omega^2)} \right] d\mathbf{r}_d, \quad (\text{C1})$$

where  $K = (q, n, l, m)$ , mode  $\mathbf{u}_K(\mathbf{r}) = \mathbf{u}(\mathbf{r}, \omega_K)$ ,  $\omega_K$  is the eigenfrequency number  $K$ ,  $\boldsymbol{\epsilon}_K$  the deformation tensor corresponding to the mode  $\mathbf{u}_K$ ,  $R_K = \mathbf{u}_K(\mathbf{r}_r) \cdot \mathbf{v}$  is the receiver term for component  $\mathbf{v}$ ,  $S_K = (\mathbf{u}_K, \mathbf{f})$  the source term. In the following, generalized spherical harmonics expansion will be used to simplify expressions. To do so, it is useful to use the contravariant canonical component of  $\delta \mathbf{c}$

$$\delta \Lambda^{\alpha\beta\gamma\eta} = \sum_{ijkl} C_{\alpha i}^* C_{\beta j}^* C_{\gamma k}^* C_{\eta l}^* \delta c_{ijkl}, \quad (\text{C2})$$

as well as for the deformation tensor associated with mode  $\mathbf{u}_K$

$$\boldsymbol{\epsilon}_K^{\alpha\beta}(\mathbf{r}) = \sum_{ij} C_{\alpha i}^* C_{\beta j}^* \boldsymbol{\epsilon}_{K,ij}(\mathbf{r}), \quad (\text{C3})$$

where  $\mathbf{C}$  is

$$(C_{i\alpha}) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & \frac{-i}{\sqrt{2}} \end{pmatrix}. \quad (\text{C4})$$

The generalized spherical harmonic expansions of  $R_K$ ,  $S_K$  and  $\boldsymbol{\epsilon}_K^{\alpha\beta}$  are well known.

$$R_K(\mathbf{r}_r) = \sum_{\alpha=-1,1} R_k^\alpha(r_r) Y_l^{\alpha m}(\theta_r, \phi_r), \quad (\text{C5})$$

$$S_K(\mathbf{r}_s) = \sum_{\alpha=-2,2} S_k^\alpha(r_s) Y_l^{\alpha m^*}(\theta_s, \phi_s), \quad (\text{C6})$$

$$\boldsymbol{\epsilon}_K^{\alpha\beta}(\mathbf{r}_d) = E_k^{\alpha\beta}(r_d) Y_l^{\alpha+\beta m}(\theta_d, \phi_d), \quad (\text{C7})$$

where  $Y_l^{\alpha m}$  is the generalized spherical harmonics and  $k = (q, n, l)$ . The expression of  $R_k^\alpha$  and  $S_k^\alpha$  can be found in Woodhouse & Girnius (1982) and  $E_k^{\alpha\beta}$  in Tanimoto (1986). Eq. (C1) can be rewritten as

$$\delta \mathbf{u}(\mathbf{r}_r, \omega) \cdot \mathbf{v} = \sum_{kk'mm'\eta\eta'} R_k^\eta Y_l^{\eta m}(\mathbf{r}_{1r}) C_{kk'mm'}^a S_{k'}^{\eta' m'^*} Y_l^{\eta' m'^*}(\mathbf{r}_{1s}), \quad (\text{C8})$$

where  $k' = (q', n', l')$ ,  $\mathbf{r}_{1s} = (\theta_s, \phi_s)$ ,  $\mathbf{r}_{1r} = (\theta_r, \phi_r)$ .

$$C_{kk'mm'}^a = \int_0^{r_\Omega} \int_{S1} \sum_{\alpha\beta\gamma\eta\iota\kappa} E_k^{\alpha\beta^*}(r) \delta \Lambda^{\alpha\beta\gamma\eta}(r) E_{k'}^{\iota\kappa}(r) Y_l^{\alpha+\beta m}(\mathbf{r}_1) Y_l^{\iota+\kappa m'^*}(\mathbf{r}_1) r^2 dr d\mathbf{r}_1, \quad (\text{C9})$$

where the integral over the Earth sphere  $V$  has been decomposed into an integration over the radius from 0 to the Earth radius ( $r_\Omega$ )

and an integration over the unit sphere surface ( $S1$ ) and where  $\mathbf{r}_1 = (\theta, \phi)$ . Using the orthogonality of the generalized spherical harmonics, we have

$$\int_{S1} Y_l^{\alpha+\beta m}(\mathbf{r}_1) Y_l^{\iota+\kappa m'^*}(\mathbf{r}_1) d\mathbf{r}_1 = \delta_{ll'} \delta_{mm'} \delta_{(\alpha+\beta)(\iota+\kappa)}. \quad (\text{C10})$$

Using the fact that no coupling occurs between spheroidal and toroidal modes due to spherically symmetric inhomogeneities and (C10), eq. (C9) reduces to

$$C_{kk'mm'}^a = C_{qnn'l} \delta_{ll'} \delta_{mm'} \delta_{qq'}, \quad (\text{C11})$$

where

$$C_{qnn'l} = \sum_{\substack{\alpha\beta\gamma\eta\iota\kappa \\ \alpha+\beta=\iota+\kappa \\ \alpha+\beta+\gamma+\eta=0}} \int_0^{r_\Omega} E_{qnl}^{\alpha\beta^*}(r) \delta \Lambda^{\alpha\beta\gamma\eta}(r) E_{qn'l}^{\iota\kappa}(r) e_{\gamma\iota} e_{\eta\kappa} r^2 dr, \quad (\text{C12})$$

where  $e_{\alpha\beta} = \sum_i C_{i\alpha} C_{i\beta}$ . Using the summation property for the generalized spherical harmonics (Li & Tanimoto 1993; Capdeville *et al.* 2000), we have

$$\sum_{m=-l}^{m=l} Y_l^{\eta' m^*}(\theta_s, \phi_s) Y_l^{\eta m}(\theta_r, \phi_r) = e^{i\eta' \gamma_{sr}} P_l^{\eta' l}(\cos(\beta_{sr})) e^{i\eta \xi_{sr}}, \quad (\text{C13})$$

where  $P_l^{\eta' l}$  are the generalized Legendre functions. If the index  $s$  is related to the ‘source’ location and  $r$  to the ‘receiver’ one, then the angle  $-\xi_{sr}$  is the backazimuth at the receiver,  $\pi - \gamma_{sr}$  the azimuth at the source and  $\beta_{sr}$  is the angular epicentral distance.

Eq. (C1) can finally be rewritten as

$$\delta \mathbf{u}(\mathbf{r}_r, \omega) \cdot \mathbf{v} = \sum_{qnn'l} R S_{qnn'l} C_{qnn'l} g_{qnl}(\omega) f_{qnn'l}(\omega), \quad (\text{C14})$$

where

$$R S_{qnn'l} = \sum_{\eta\eta'} R_{qnl}^\eta S_{qn'l}^{\eta'} e^{i\eta' \gamma_{sr}} P_l^{\eta' l}(\cos(\beta_{sr})) e^{i\eta \xi_{sr}}, \quad (\text{C15})$$

$$g_{qnl}(\omega) = \frac{1}{\omega_{nl}^2 - \omega^2} \quad (\text{C16})$$

and

$$f_{qnl}(\omega) = \frac{1}{i\omega(\omega_{nl}^2 - \omega^2)}. \quad (\text{C17})$$

Going back from the frequency domain to the time domain can be done analytically using the residue theorem or with a fast Fourier transform. It can be seen from the final expression (C14) that the computation only involves coupling along the radial order  $n$ , which reduces drastically the amount of computation compared to the classical mode coupling due to a 3-D inhomogeneities. From (C14), for a given earth model and a given parametrization (leading to a set of  $\delta \Lambda$ ), the partial derivative matrix  $\mathbf{G}^i$  can easily be built.