

# Surface wave tomography: global membrane waves and adjoint methods

D. Peter,<sup>1</sup> C. Tape,<sup>2</sup> L. Boschi<sup>1</sup> and J. H. Woodhouse<sup>3</sup>

<sup>1</sup>Institute of Geophysics, ETH Zurich, Höggerberg HPP, CH-8093 Zurich, Switzerland. E-mail: dpeter@erdw.ethz.ch

<sup>2</sup>Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

<sup>3</sup>University of Oxford, Department of Earth Sciences, Parks Road, Oxford OX1 3PR, UK

Accepted 2007 July 23. Received 2007 June 15; in original form 2007 August 18

## SUMMARY

We implement the wave equation on a spherical membrane, with a finite-difference algorithm that accounts for finite-frequency effects in the smooth-Earth approximation, and use the resulting ‘membrane waves’ as an analogue for surface wave propagation in the Earth. In this formulation, we derive fully numerical 2-D sensitivity kernels for phase anomaly measurements, and employ them in a preliminary tomographic application. To speed up the computation of kernels, so that it is practical to formulate the inverse problem also with respect to a laterally heterogeneous starting model, we calculate them via the adjoint method, based on backpropagation, and parallelize our software on a Linux cluster. Our method is a step forward from ray theory, as it surpasses the inherent infinite-frequency approximation. It differs from analytical Born theory in that it does not involve a far-field approximation, and accounts, in principle, for non-linear effects like multiple scattering and wave front healing. It is much cheaper than the more accurate, fully 3-D numerical solution of the Earth’s equations of motion, which has not yet been applied to large-scale tomography. Our tomographic results and trade-off analysis are compatible with those found in the ray- and analytical-Born-theory approaches.

**Key words:** adjoint methods, Born approximation, membrane waves, scattering theory, seismic tomography, surface waves.

## 1 INTRODUCTION

One of the important challenges in seismology is to enhance the tomographic resolution of the Earth’s interior. A way to achieve this goal is by elaborating more accurate theoretical descriptions of seismic wave propagation and using them to formulate the tomographic inverse problem. The Born approximation (single-scattering theory) represents a possibly significant improvement with respect to simple ray theory; for some time now it has been known to be valid at least for weak scattering in the Earth (Hudson & Heritage 1981; Wu & Aki 1985). On the basis of the Born approximation, the ‘banana–doughnut’ paradox, or the prediction that the sensitivity of seismic traveltimes observations be maximum away from the ray-theoretical path, was pointed out most clearly by Marquering *et al.* (1999) (but see also Kennett 1972; Woodhouse & Gernius 1982; Snieder 1987; Snieder & Nolet 1987; Li & Tanimoto 1993; Li & Romanowicz 1995); numerous studies have been made to understand the nature of this phenomenon (Dahlen *et al.* 2000; Hung *et al.* 2000; Zhao *et al.* 2000; Spetzler & Snieder 2001; Spetzler *et al.* 2002; Baig *et al.* 2003; Tanimoto 2003; Baig & Dahlen 2004; de Hoop *et al.* 2005; Tromp *et al.* 2005).

The Born approximation, while representing an improvement with respect to simple ray theory, cannot reproduce non-linear effects that take place in reality (Wielandt 1987; Nolet & Dahlen 2000; Hung *et al.* 2001; Baig *et al.* 2003). In addition, implementations of

Born theory found in the literature are typically based on far-field expressions of the wavefield, giving rise to singularities in the computed Fréchet derivatives (sensitivity kernels) (Favier *et al.* 2004). In practice, the forward problem is formulated under the assumption that scattering occurs only at distances from source and receiver much larger than the wavelength, but its solution is then used to compute the values of sensitivity kernels in the vicinity of source and receiver as well, where the kernels are actually singular. This singularity appears to be integrable (Friederich 1999, Appendix E), but it remains unclear to what extent these asymptotic kernels are valid near the source and receiver. Capdeville (2005) proposed an efficient way to overcome this problem in the computation of sensitivity kernels, combining adjoint methods (Tromp *et al.* 2005, and references therein) and normal mode summation, though his approach has not yet been applied in practice to the inversion of real seismic data.

Direct numerical integration of the equations of motion is another way to avoid the far-field approximation, and simulate, at least to some extent, the non-linear phenomena mentioned above. Computational power has grown tremendously in recent years thanks to a combined improvement in processor performance and in size of multi-processors clusters (Bunge & Tromp 2003; Boschi *et al.* 2007); over the last decade, the average growth-per-year in the computational performance of supercomputers has slightly exceeded Moore’s law, which predicts a factor of 2 in 18 months (Strohmaier

*et al.* 2005). Seismologists with access to large parallel computers are now able to calculate seismic waves propagating through a 3-D, complex medium closely resembling the Earth. Nevertheless, such simulations can take quite a long time even for a single earthquake and on large cluster systems; for example, Tsuboi *et al.*'s (2002) simulation of the Denali earthquake in a very realistic earth model took about 15 hr on one of the fastest supercomputers in the world. It is therefore, too time-consuming to perform the large number of full numerical integrations required to set up an inverse problem numerically unless one simplifies the problem (e.g. Tape *et al.* 2007).

Here we explain how the 3-D problem of fundamental-mode surface wave propagation is represented by a 2-D, zero-thickness membrane analogue. We then use the membrane wave analogue to compute sensitivity functions relating surface wave phase measurements to lateral anomalies in the phase velocity of surface waves. This is a notoriously time-consuming endeavour, but we reduce its cost through the application of an adjoint-method approach (e.g. Tarantola 1984) similar to that of Tromp *et al.* (2005). We apply the sensitivity functions in an inversion of real data to derive phase-velocity maps of the Earth, which we compare in the last part of this work with those obtained from ray-theory and analytical Born-theory tomography (e.g. Spetzler *et al.* 2002; Boschi 2006).

## 2 THE FORWARD PROBLEM: MEMBRANE WAVES

Membrane waves as an analogue for surface waves were introduced by Tanimoto (1990), to investigate locally the strong effects of lateral heterogeneity on short-period surface waves (10 and 20 s). We follow the same approach to derive a numerical model for the propagation of intermediate-period surface waves at the global scale, which requires special spherical grids and numerical techniques (e.g. Sword *et al.* 1986). In the interest of simplicity, we give only a detailed theoretical treatment of the Love-wave case below. Working with a Rayleigh wave *Ansatz* leads to an algebraically more complicated, but qualitatively analogous derivation (Tanimoto 1990; Tromp & Dahlen 1993).

### 2.1 Theory

Given the equation of motion (Aki & Richards 2002)

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

with  $\rho$  denoting density,  $\mathbf{u}$  displacement,  $\boldsymbol{\tau}$  stress, and accompanied by initial conditions accounting for the seismic source, we Fourier-transform it and replace  $\mathbf{u}$  with the Love wave *Ansatz* (Tanimoto 1990)

$$\mathbf{u}_L = W(r)(-\hat{\mathbf{r}} \times \nabla_1)s(\vartheta, \varphi), \quad (2)$$

with  $W(r)$  depending only upon depth, and  $\hat{\mathbf{r}}$  denoting the unit vector,  $\nabla_1$  the surface gradient (e.g. Dahlen & Tromp 1998) and  $s(\vartheta, \varphi)$  a scalar 'potential' at colatitude  $\vartheta$  and longitude  $\varphi$ . It follows that for an isotropic stress-strain relation and in the smooth-Earth approximation,  $s$  satisfies

$$\nabla_1^2 s + \frac{\omega^2}{c^2(\vartheta, \varphi; \omega)} s = 0, \quad (3)$$

valid at each frequency  $\omega$  and corresponding phase velocity  $c$  (Tanimoto 1990; Tromp & Dahlen 1993; Yoshizawa & Kennett 2005). We solve eq. (3) numerically in the time domain, with a

prescribed initial displacement and source-time function. Note that we have implicitly assumed the medium to be isotropic, with smooth lateral variations of the shear modulus, so that its gradient can be neglected; in practice, we are restricting ourselves to a similar approximation as those implicit in perturbation (Born) theories based upon ray theory (e.g. Dahlen *et al.* 2000; Spetzler *et al.* 2002).

In order for an analytical solution to be available in the homogeneous Earth case, we prescribe a forcing term

$$f(\vartheta, \varphi) = g(\Delta) h(t), \quad (4)$$

consisting of an initial displacement

$$g(\Delta) = \frac{e^{-\Delta^2/2\mu^2}}{\mu^2}, \quad (5)$$

with  $\Delta(\vartheta, \varphi) \in [0, \pi]$  denoting arc-distance from the source, and a source-time function

$$h(t) = \frac{-t}{\sigma^2} \frac{e^{-t^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \quad (6)$$

(Tape 2003, eq. 3.31). Here  $\sigma$  and  $\mu$  act as source parameters, governing the characteristic frequency content of the source. To avoid the introduction of absorbing boundary conditions and to allow for multi-orbit surface wave propagation, the membrane spans the whole globe.

### 2.2 Meshing the Earth's surface

It is impossible to evenly distribute more than 20 points on the surface of the sphere, a distribution of points describing a platonic dodecahedron. Modelling on the sphere at regional scale lengths requires thousands of points, and thus the grids will have at least some undesired irregularities. Spherical geodesic grids of triangular faces, first introduced in the context of meteorological flow modelling (Sadourny *et al.* 1968; Williamson 1968), are one approach to the non-trivial problem of uniformly discretizing the Earth's surface (Cui & Freedman 1997; Saff & Kuijlaars 1997).

In practice, the sphere is first discretized by a platonic solid, that is, a regular polyhedron consisting of identical cell faces. We find the initial triangular mesh combining the vertices of a dodecahedron and an icosahedron (Tape 2003), also known as a truncated icosahedron or a buckyball. We then refine the grid iteratively, using the midpoints of the three sides of each spherical triangle to divide it into four new smaller triangles, forming the next finer grid. This method will be referred as *dyadic refinement* (Baumgardner & Frederickson 1985). The corresponding hexagonal/pentagonal (in the following loosely referred to as hexagonal) grid is constructed from the triangular grid. The corner of each hexagonal or pentagonal face is found by calculating the central point of each spherical triangle from the corresponding triangular grid. Vertices of the triangular grids then coincide with the cell centres of the hexagonal grids.

As we shall show in the next section,  $s$  is evaluated at the vertices of the triangular grid, while the computation of its Laplacian involves the areas of the hexagonal grid cells, so that the properties of both grids are relevant to our implementation. Comparisons in Table 1 show that our hexagonal grid has a more uniform distribution than others based on the icosahedron as the initial triangular grid (e.g. Heikes & Randall 1995a; Wang & Dahlen 1995). A further improvement of the distribution could still lead to an even more accurate numerical solution; in the next section we show that our hexagonal grid is sufficiently good to provide a valuable basis for the numerical computations.

Table 1. Properties of spherical grids. Comparison between spherical grids based on icosahedral (ico) and icosahedral-dodecahedral (ico-dod) initial platonic solids. The membrane wave calculations use the hexagonal grid based on the second combination of two platonic solids (ico-dod). Values for the icosahedral case are taken from the twisted grid used by Heikes & Randall (1995a).

Refinement level		Ratio of cell area $A_{\min}/A_{\max}$	Ratio of cell distance $d_{\min}/d_{\max}$
0	ico	0.885	0.881
	ico-dod	0.941	0.894
1	ico	0.774	0.848
	ico-dod	0.914	0.861
2	ico	0.763	0.840
	ico-dod	0.907	0.852
3	ico	0.742	0.838
	ico-dod	0.878	0.850
4	ico	0.736	0.837
	ico-dod	0.870	0.849
5	ico	0.733	0.837
	ico-dod	0.868	0.849
6	ico	— <sup>a</sup>	— <sup>a</sup>
	ico-dod	0.868	0.849

<sup>a</sup>No information.

### 2.3 Finite-difference scheme

Both global (e.g. spectral) and local (e.g. finite-difference) numerical methods have been applied to meteorological problems associated with spherical geodesic grids (see Stuhne & Peltier 1996; Randall *et al.* 2002). Inspired by the studies of Heikes & Randall (1995a,b), we use a simple finite-difference scheme to compute the Laplacian, which is required to solve eq. (3) for the wavefield  $s$ . The Laplacian of a function  $\alpha(\vartheta, \varphi)$  at a cell centre point  $\mathbf{P}_i$  on the hexagonal grid described above is approximated by

$$\nabla^2 \alpha|_{\mathbf{P}_i} \approx \frac{1}{A_i} \sum_{n=1}^N \frac{l_n}{L_n} (\alpha_n - \alpha_i), \quad (7)$$

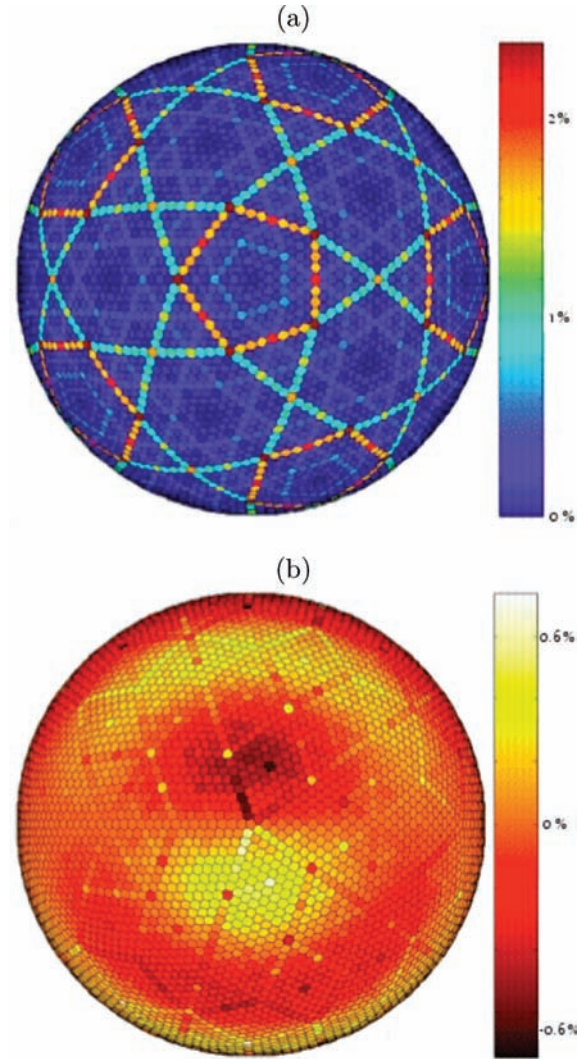
where  $N$  denotes the number of edges of the  $i$ th cell (5 or 6),  $l_n$  is the length of its  $n$ th edge and  $L_n$  the distance between the centres of the  $i$ th cell and of its neighbour, which shares its  $n$ th edge.  $\alpha_i$  denotes the value of  $\alpha$  at  $\mathbf{P}_i$  and  $\alpha_n$  the value of  $\alpha$  at the centres of the neighbouring cells.  $A_i$  is the area of the spherical polygon identifying the cell on the sphere's surface. All these parameters are computed numerically at the outset.

Eq. (7) shows that the calculation of the Laplacian for a given cell requires only information on the neighbouring cells. This is why this scheme is said to be a ‘local’ method, in contrast to other ‘global’ numerical implementations like the spectral methods mentioned above. The advantage of local methods is that they can be parallelized in a very efficient way.

We validate eq. (7) using spherical harmonic functions  $Y_{lm}$  (Dahlen & Tromp 1998, eq. B.11), which are eigenfunctions of the Laplacian operator,

$$\nabla^2 Y_{lm}(\vartheta, \varphi) = -l(l+1)Y_{lm}(\vartheta, \varphi). \quad (8)$$

Hence, differences between the numerically evaluated Laplacian of  $Y_{lm}$ , and its analytical values, are readily calculated at the location of each cell. We find that differences are not uniformly distributed on the spherical grid; in fact, although errors become very small as soon as we have a sufficiently fine grid, there exist certain cells where, above a certain level, further refinements of the mesh do not reduce the error significantly.

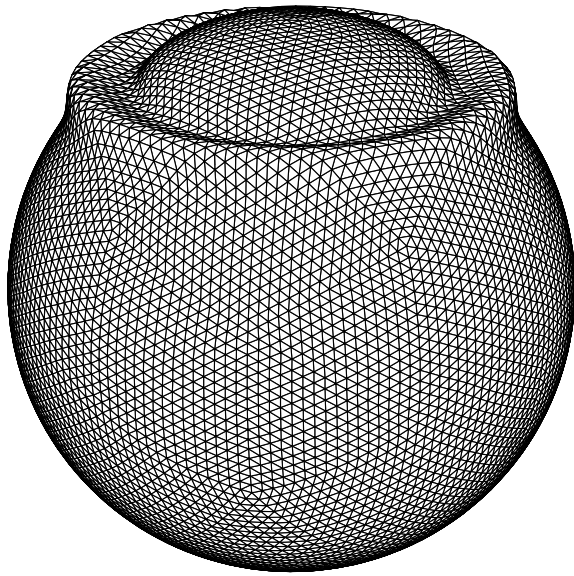


**Figure 1.** Properties of the hexagonal grid. (a) Distortion measured by the midpoint ratio of distance between cell centres and cell edges to corresponding cell edge length (Heikes & Randall 1995b). (b) Accuracy of Laplacian for a spherical harmonic function ( $l=6, m=1$ ). Differences between numerical and analytical values are normalized to the maximum value of the analytical Laplacian. View from North Pole down projected onto the equatorial plane.

We explain this effect introducing ‘grid distortion’ as the distance between the midpoint of the segment connecting two neighbouring cell centres, and the midpoint of the corresponding cell edge, divided by the cell edge length (Heikes & Randall 1995b, section 2). At each cell, we average grid distortion over all its neighbours, and plot the result in Fig. 1(a). When the plotted value is zero, the cell is symmetric and the grid is not distorted at the corresponding location. We find the Laplacian to be nearly second-order accurate at most gridpoints (see also Heikes & Randall 1995b). A few cells exist with a particularly high distortion ( $\sim 2$  per cent or less). Table 1 confirms, nevertheless, that our grid is more uniform (hence, less distorted) than alternatives found in current literature.

In a numerical simulation relying on eq. (7), cells with high distortion will generate non-physical scattered waves. We test the accuracy of the numerical scheme comparing numerical and analytical solutions found for a constant-velocity membrane. Using the source described by eqs (5) and (6), the following analytical solution was





**Figure 2.** Snapshot of wave propagation on the spherical membrane for a source at the North Pole. Wavefield  $s$  is plotted on a spherical-triangular grid with 7682 vertices.

determined for example by Tape (2003, eq. 3.34):

$$s(\vartheta, \varphi, t) = c^2 \sum_{l=0}^{\infty} \left( l + \frac{1}{2} \right) I_l(\mu) \cos(\omega_l t) e^{-\omega_l^2 \sigma^2 / 2} P_l[\cos(\vartheta)], \quad (9)$$

where

$$\omega_l = \frac{c \sqrt{l(l+1)}}{a} \quad (10)$$

with  $a$  the radius of the sphere, and the integrals

$$I_l(\mu) = \int_0^\pi P_l[\cos(\theta)] \frac{e^{-\theta^2 / 2\mu^2}}{\mu^2} \sin(\theta) d\theta \quad (11)$$

are calculated numerically with the Legendre polynomials  $P_l$  of angular degree  $l$ .

Fig. 2 shows a snapshot of the scalar potential solution to the wave eq. (3) derived numerically over the whole sphere. Fig. 3 shows the resulting scalar potential solutions and phaseshifts between the analytical and numerical solutions for Love waves at 150 s period. Phaseshifts are shown for filtered and unfiltered solutions; the simple source mechanism described by eqs (5) and (6) excites a wider range of frequencies, which have to be filtered out in order to isolate the frequency of interest. We examine this in more detail in Section 3.3.2

We find that differences between the analytical and numerical solutions are small enough that the numerical algorithm can be considered valid, and can be applied to evaluate the effects of small lateral heterogeneities, or formulate a tomographic inverse problem. We conduct numerical integrations on the grid defined by six successive dyadic refinements, yielding 122 882 hexagonal or pentagonal cells. On the basis of Fig. 3, we consider this mesh a good compromise between the cost of further refinements, and inaccuracies introduced by making the grid coarser. The grid spacing, or average distance between the centres of neighbouring cells, is  $\sim 70$  km, corresponding to about 10 gridpoints per dominant wavelength for the reference case of 150 s period waves.

A final remark concerns the stability of our numerical scheme. If the time step  $dt$  becomes too small, the numerical errors can grow

exponentially and become unstable within the time window of the simulation. On the other hand,  $dt$  must satisfy

$$dt \leq dx/R, \quad (12)$$

where  $dx$  is the average distance between grid nodes, and the threshold value  $R$  has to be determined as described, for example, by Moczo *et al.* (2004). In the case of our spherical grid, Tape (2003) found empirically

$$dt \approx \frac{dx}{\bar{c}\sqrt{2}}, \quad (13)$$

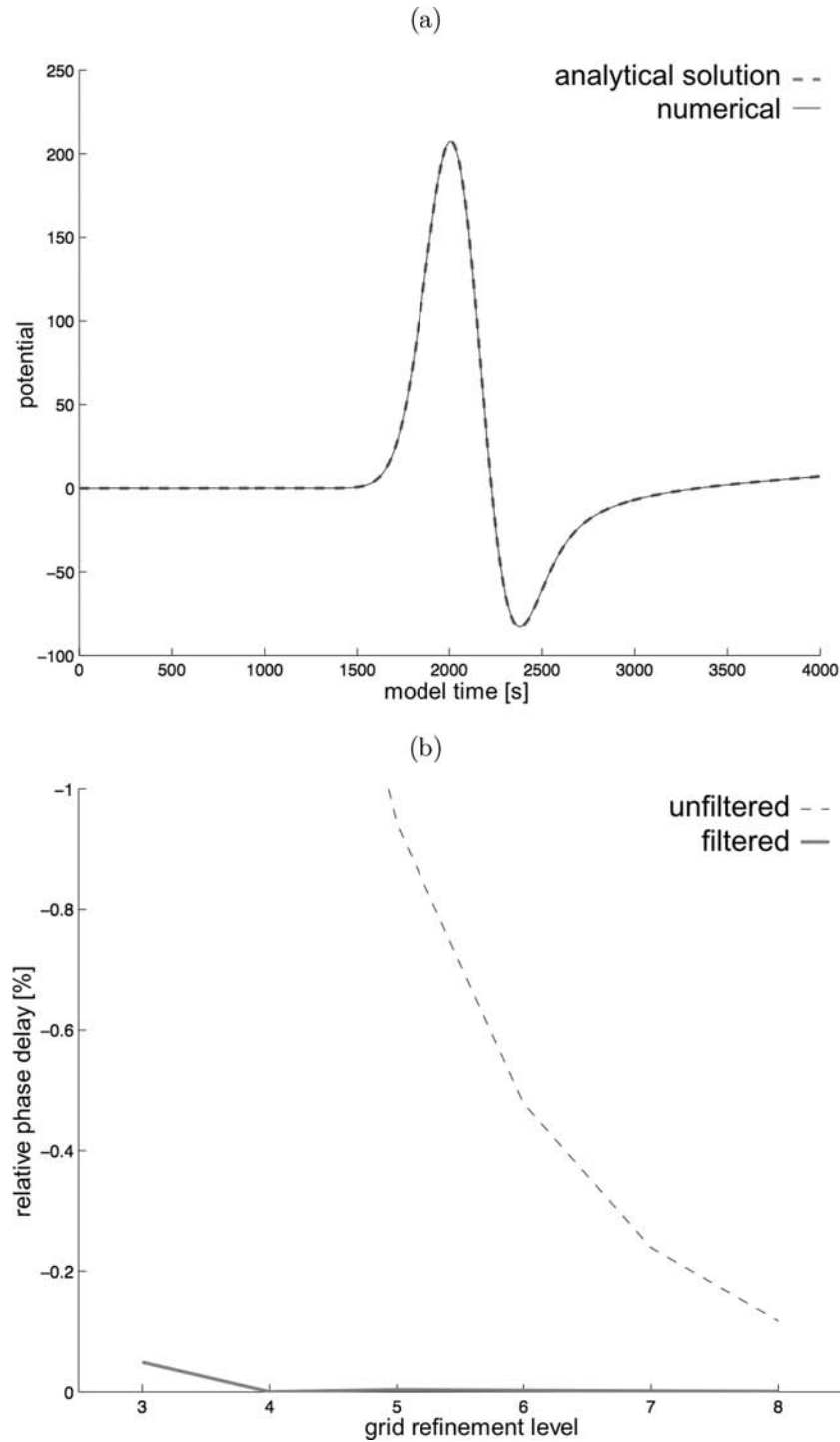
where  $\bar{c}$  is the average phase velocity of the model. Simulations with a time step smaller than (13) will require a longer computation time without increasing the precision. The time step also limits the sampling rate of modelled waves; in Section 4, we will apply a quadratic interpolation between sampled times to cross-correlate accurately our computed traces (e.g. Smith & Serra 1987; Press *et al.* 1992).

## 2.4 Computational considerations

The main advantage of collapsing a 3-D problem to two dimensions is an order-of-magnitude gain in the speed of simulations. Our software is parallelized to optimize its performance on a Linux cluster. We use the implementation MPICH of the standard message passing interface on a 16-processor cluster. Software performance on parallel computers depends on the amount of communication between different processes needed during a simulation. If no communication were needed, the calculation time would decrease linearly as a function of the number of available processors. The more communication required, the slower the performance will get, and, above a certain number of processors, no further gain in speed will be possible.

Table 2(a) shows how the performance of our implementation scales with respect to the number of processors. It can be seen that if the latter is doubled, runtime decreases by a factor of  $\sim 2$ . Above eight processors, the gain in speed falls off. Note that in principle we should expect the simulation time to increase by a factor 8 for a next finer grid, as there will be four times more grid cells due to the dyadic refinement, and the number of time steps will double due to a halving of the space step. Table 2(b) shows the runtimes for single simulations conducted with different grid spacings employing 16 processors in parallel. A single simulation, running in parallel on clusters larger than ours, will nevertheless be punished by the need of data communication between the processes. Since one simulation can be completed within a minute on a single processor (for grid spacings around 70 km), a large cluster system has the obvious advantage that each processor can run a separate simulation.

A single simulation of surface wave propagation via the full numerical integration of the equation of motion in 3-D (e.g. spectral-element methods) would take minutes even on large cluster systems, and can be carried out in seconds by our approximate, 2-D membrane wave algorithm on a small cluster. While analytical methods like that of Spetzler *et al.* (2002) are only proved to perform well in a spherical-background-Earth scenario, we show below that ours, through an application of the adjoint method (Tarantola 1984), can provide a full kernel library in a heterogeneous reference model as well, in a reasonable amount of time.



**Figure 3.** Accuracy of numerical membrane waves. (a) Analytically (dashed) and numerically (solid) calculated membrane wave  $s$ , generated from a source at  $(0^\circ\text{N}, 0^\circ\text{E})$  and recorded at  $(0^\circ\text{N}, 90^\circ\text{E})$ . (b) Phase-shift between the analytical and numerical, filtered (solid) and unfiltered (dashed) solutions shown in (a), for different levels of grid refinements. Filtering is discussed in Section 3.3.2

### 3 MEMBRANE-WAVE SENSITIVITY FUNCTIONS

#### 3.1 A ‘direct’ approach

Let us introduce the single-scattering sensitivity function  $K(\vartheta, \varphi)$  (e.g. Spetzler *et al.* 2002), relating a local relative perturbation in phase velocity  $\delta c/c$  on the sphere’s surface  $\Omega$  to a relative phase

anomaly  $\delta\Phi/\Phi$ , via the linear equation

$$\frac{\delta\Phi}{\Phi} = \int_{\Omega} K(\vartheta, \varphi) \frac{\delta c}{c}(\vartheta, \varphi) d\Omega. \quad (14)$$

It is intended that  $K(\vartheta, \varphi)$  depends on the source-station distance, or, in the case of a laterally heterogeneous Earth (membrane), on the locations of source and station.

Table 2. Performance efficiency. Runtimes with (a) different numbers of processors and an average grid spacing of 70 km, (b) 16 processors<sup>a</sup>, different grid spacings.

Number of processors	Runtime (s)
(a)	
1	60.9
2	31.8
4	15.9
8	8.5
16	5.5
Grid spacing (km)	Runtime (s)
(b)	
278	0.1
139	0.7
70	5.5
35	42.0
17	340.3

<sup>a</sup>Processor type: AMD Opteron 64-bit, 2 GHz clock speed.

For any given source-station combination,  $K(\vartheta, \varphi)$  can be determined numerically, performing a set of simulations on a given background phase-velocity model  $c(\vartheta, \varphi)$ , with one small perturbation at a single cell centred at  $(\vartheta_i, \varphi_i)$ ,

$$\frac{\delta c}{c} = f_i(\vartheta, \varphi) = \begin{cases} \gamma & \text{in the gridcell centred at } (\vartheta_i, \varphi_i) \\ 0 & \text{everywhere else.} \end{cases} \quad (15)$$

Eq. (14) then becomes

$$\frac{\delta \Phi_i}{\Phi} = \int_{\Omega} K(\vartheta_i, \varphi_i) \gamma \, d\Omega = K(\vartheta_i, \varphi_i) \gamma A_i, \quad (16)$$

where  $A_i$  is the area of the  $i$ th cell. We denote  $\delta \Phi_i$  the phaseshift associated with the perturbation (15), which we calculate by a membrane simulation and subsequent cross-correlation of the modelled phase with the unperturbed one. Then, from (16),

$$K(\vartheta_i, \varphi_i) = \frac{\delta \Phi_i}{\Phi A_i \gamma}, \quad (17)$$

and  $K(\vartheta_i, \varphi_i)$  can be found from the numerically calculated  $\delta \Phi_i$ . In our calculations of  $\delta \Phi_i$ , we compute cross-correlation on the basis of Press *et al.*'s (1992) routine, but with quadratic subsample precision (Smith & Serra 1987).

Note that numerical integration of eq. (3) on the membrane involves a simplified representation of the source in terms of initial displacement and source-time function (Section 2.1). We must therefore, neglect the effects on wave propagation of specific seismic source mechanisms (Zhou *et al.* 2004; Yoshizawa & Kennett 2005). This will be the subject of further investigations.

In principle, calculating  $K(\vartheta_i, \varphi_i)$  for each cell  $i$  requires as many simulations as there are cells. This exercise then has to be repeated, in the case of a homogeneous reference Earth [constant  $c(\vartheta, \varphi)$ ], for a dense set of epicentral distances spanning the range of true epicentral distances at which observations are available. In the case of a laterally heterogeneous reference Earth, it has to be repeated for each combination of source and station locations for which observations are available. The latter endeavour is too

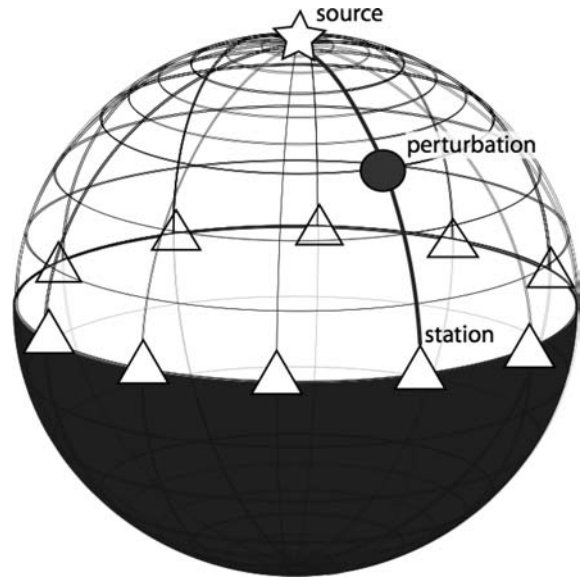


Figure 4. Set-up for the ‘direct’ approach to computing sensitivity kernels (Section 3.1.1), with one source (star) located at the North Pole, the scatterer position (circle) varying along a single meridian, and a set of receiver stations (triangles) located along parallels.

costly and we have discarded it; however, in the case of a constant reference phase velocity, it is feasible to calculate kernels via eq. (17).

### 3.1.1 Geometrical setup

We reduce the number of simulations needed to find  $K(\vartheta, \varphi)$  in the ‘direct’ approach as follows (Tong *et al.* 1998): after placing the source at, say, the North Pole, we perform one simulation for each scatterer location along one chosen meridian (Fig. 4). For each simulation, we save modelled traces at an array of receivers located along the parallels, spaced  $1^\circ$  in longitude and latitude from each other. If the background earth model is homogeneous, what matters is not the absolute locations of scatterer and receiver, but their relative positions. From the set-up described above, we therefore, find the same traces that we would find considering one receiver at a time, and performing one simulation per possible scatterer location.

In practice, this results in reducing the number of required simulations from 64 442 to 181 (for the reference case of  $1^\circ \times 1^\circ$  sampling of  $K(\vartheta, \varphi)$ , and using 360 receivers simultaneously). Running our parallel algorithm on 16 processors, with a gridspacing of about 70 km (level six hexagonal grid), we entirely determine  $K(\vartheta, \varphi)$ , for a given epicentral distance, in about 6 min. Without the simplification introduced here, the same computation would take approximately 3 d.

### 3.1.2 Non-linearity

In principle, sensitivity functions derived as in Section 3.1 should not depend on the value of the imposed phase-velocity perturbation  $\gamma$ . However, as  $\gamma$  grows, the problem eventually becomes non-linear, and the linearized eq. (14) ceases to be valid, together

with the concept itself of sensitivity kernels. Testing the stability of our algorithm, we have found that, above a certain threshold ( $\gamma \sim 2$  per cent), the mentioned non-linearity comes into play and numerical kernels we find are slightly but visibly affected by changes in the value of  $\gamma$ .

Hung *et al.* (2001) found a similar effect from a set of 3-D spectral element simulations. In their fig. 17 they plot traveltime anomaly found by cross-correlation (equivalent to our  $\delta\Phi$ ) as a function of imposed heterogeneity ( $\gamma$  in Section 3.1 here,  $\varepsilon$  in Hung *et al.*'s (2001) notation). While Born theory requires that a linear relationship exists between  $\delta\Phi$  and  $\gamma$ , the corresponding curve obtained from numerical results is a straight line only for small values of  $\gamma$ . Hung *et al.* (2001, section 5.6) find an asymmetry in the dependence of  $\delta\Phi$  on  $\gamma$  for positive versus negative values of  $\gamma$ . They explain this result as a combination of wave front healing and the 'over-healing' effect of the cross-correlation technique. In our case there is not such a strong asymmetry between the effect of negative and positive anomalies.

### 3.2 The adjoint approach

Our 'membrane' algorithm is efficient enough that, on a homogeneous reference Earth,  $K(\vartheta, \varphi)$  can be computed by a large set of direct simulations. The number of simulations required to determine  $K(\vartheta, \varphi)$ , however, is much larger when the reference Earth is laterally heterogeneous, and the 'direct' approach outlined above ceases to be practical. Tromp *et al.* (2005) give an overview of the application of backpropagation to the calculation of sensitivity functions, resulting in the 'adjoint methods' introduced by Tarantola (1984) or Talagrand & Courtier (1987). In this approach, regardless of the complexity of the reference model,  $K(\vartheta, \varphi)$  for a given source–receiver pair can be fully determined with two simulations only: one for the forward-propagating wavefield, from the source to the receiver; another for the backpropagating wavefield, from the receiver to the source. At each point  $(\vartheta, \varphi)$ ,  $K(\vartheta, \varphi)$  is found algebraically as a function of the forward- and backpropagating wavefields.

Next we provide a formulation of the adjoint method for the case of surface wave phase-anomaly observations, to be inverted tomographically in our spherical membrane approach. Part of our treatment is very similar to that of Yoshizawa & Kennett (2005, section 2), except that we prefer to work in the time domain; the rest follows Tromp *et al.* (2005, sections 2 and 4.1).

#### 3.2.1 Writing displacement anomalies in terms of the Green's function

The time-domain version of eq. (3) is given by

$$\left[ \frac{1}{c^2(\mathbf{x})} \partial_t^2 - \nabla_1^2 \right] s(\mathbf{x}, t) = f(\mathbf{x}_s, t), \quad (18)$$

with  $\mathbf{x} = (\vartheta, \varphi)$  for brevity,  $\mathbf{x}_s$  denoting the location of the source, and  $f(\mathbf{x}_s, t)$  a forcing term representing the source. Let us symbolize  $c_0$  and  $s_0$ , respectively, the values of phase velocity and of the solution to (18) associated with our reference earth (membrane) model [ $c_0(\mathbf{x})$  need not be uniform]; then

$$\left[ \frac{1}{c_0^2(\mathbf{x})} \partial_t^2 - \nabla_1^2 \right] s_0(\mathbf{x}, t) = f(\mathbf{x}_s, t). \quad (19)$$

After adding a term  $\frac{1}{c_0^2(\mathbf{x})} \partial_t^2 s(\mathbf{x}, t)$  to both sides of (18),

$$\left[ \frac{1}{c_0^2(\mathbf{x})} \partial_t^2 - \nabla_1^2 \right] s(\mathbf{x}, t) = \frac{1}{c_0^2(\mathbf{x})} \left[ 1 - \frac{c_0^2(\mathbf{x})}{c^2(\mathbf{x})} \right] \partial_t^2 s(\mathbf{x}, t) + f(\mathbf{x}_s, t). \quad (20)$$

Introducing

$$\delta c(\mathbf{x}) = c(\mathbf{x}) - c_0(\mathbf{x}), \quad (21)$$

we simplify the right-hand side of (20) noting that

$$1 - \frac{c_0^2(\mathbf{x})}{c^2(\mathbf{x})} = \frac{2c(\mathbf{x})\delta c(\mathbf{x}) - [\delta c(\mathbf{x})]^2}{c^2(\mathbf{x})} \approx \frac{2\delta c(\mathbf{x})}{c(\mathbf{x})}, \quad (22)$$

after neglecting perturbations of second order in  $\delta c$ . Likewise, introducing  $\delta s(\mathbf{x}, t) = s(\mathbf{x}, t) - s_0(\mathbf{x}, t)$  and using the relation (19), we reduce (20) to

$$\left[ \frac{1}{c_0^2(\mathbf{x})} \partial_t^2 - \nabla_1^2 \right] \delta s(\mathbf{x}, t) = \frac{2\delta c(\mathbf{x})}{c_0^2(\mathbf{x})c(\mathbf{x})} \partial_t^2 s(\mathbf{x}, t). \quad (23)$$

Since we are neglecting terms of second order in  $\delta c$ ,  $s(\mathbf{x}, t)$  can be replaced by  $s_0(\mathbf{x}, t)$  in the right-hand side.

If we know the Green's function (e.g. Dahlen & Tromp 1998) for the membrane problem in question, defined as the solution  $G(\mathbf{x}, \mathbf{x}'; t, t')$  to

$$\left[ \frac{1}{c_0^2(\mathbf{x})} \partial_t^2 - \nabla_1^2 \right] G(\mathbf{x}, \mathbf{x}'; t, t') = -\delta(\mathbf{x} - \mathbf{x}')\delta(t - t'), \quad (24)$$

we can use it to find a solution to (23), that is,

$$\begin{aligned} \delta s(\mathbf{x}, t) = & \int_0^t \int_{\Omega} -\frac{2}{c_0^2(\mathbf{x}')} G(\mathbf{x}, \mathbf{x}'; t - t') \\ & \times \partial_t^2 s_0(\mathbf{x}', t') \frac{\delta c(\mathbf{x}')}{c(\mathbf{x}')} d\mathbf{x}' dt'. \end{aligned} \quad (25)$$

Note that the procedure we have followed so far is analogous, for example, to those of Snieder & Nolet (1987), Dahlen *et al.* (2000) and Boschi (2006).

#### 3.2.2 Writing traveltime (or phase) anomalies in terms of the Green's function

As in Dahlen *et al.* (2000), we assume that the phase anomaly between two seismograms can be defined as the value of phase-shift between the two, that maximizes their cross-correlation. In terms of traveltime  $T$ , and relative perturbations in traveltime  $\frac{\delta T}{T_0}$  (which coincide with relative perturbations in phase), this implies

$$\frac{\delta T}{T_0} = \frac{1}{T_0} \frac{1}{N} \int_0^T w(t) \partial_t s(\mathbf{x}_r, t) \delta s(\mathbf{x}_r, t) dt \quad (26)$$

(Tromp *et al.* 2005, eq. 41; Marquering *et al.* 1999), where  $T$  is the duration of the seismogram and  $N$  is a normalization factor given by

$$N = \int_0^T w(t) s(\mathbf{x}_r, t) \partial_t^2 s(\mathbf{x}_r, t) dt, \quad (27)$$

where  $w(t)$  denotes the cross-correlation window and  $\mathbf{x}_r$  the location of the receiver. Note that  $N$  will be different for each source–station pair. Substituting (25) into (26),

$$\begin{aligned} \frac{\delta T}{T_0} = & \frac{1}{T_0} \frac{1}{N} \int_0^T w(t) \partial_t s(\mathbf{x}_r, t) \int_0^t \int_{\Omega} -\frac{2}{c_0^2(\mathbf{x}')} \\ & \times G(\mathbf{x}_r, \mathbf{x}'; t - t') \partial_t^2 s(\mathbf{x}', t') \left[ \frac{\delta c(\mathbf{x}')}{c(\mathbf{x}')} \right] d\mathbf{x}' dt' dt. \end{aligned} \quad (28)$$



### 3.2.3 Finding membrane sensitivity kernels in the adjoint approach

It is now convenient to make use of the equality (Tromp *et al.* 2005)

$$\int_0^T h(t) \int_0^t g(t') dt' dt = \int_0^T g(t) \int_0^{T-t} h(T-t') dt' dt, \quad (29)$$

valid for any integrable functions  $h(t)$ ,  $g(t)$ , to rewrite eq. (28) in the form

$$\begin{aligned} \frac{\delta T}{T_0} = & \int_{\Omega} \frac{1}{T_0} \frac{1}{N} \int_0^T -\frac{2}{c_0^2(\mathbf{x}')} \partial_t^2 s(\mathbf{x}', t) \frac{\delta c(\mathbf{x}')}{c(\mathbf{x}')} \\ & \times \int_0^{T-t} G(\mathbf{x}', \mathbf{x}_r, T-t-t') \\ & \times w(T-t') \partial_t s(\mathbf{x}_r, T-t') dt' d\mathbf{x}'. \end{aligned} \quad (30)$$

If one then denotes

$$\begin{aligned} \bar{s}^\dagger(\mathbf{x}', \mathbf{x}_r, T-t) = & \frac{1}{N} \int_0^{T-t} G(\mathbf{x}', \mathbf{x}_r, T-t-t') \\ & \times w(T-t') \partial_t s(\mathbf{x}_r, T-t') dt', \end{aligned} \quad (31)$$

eq. (30) is reduced to

$$\begin{aligned} \frac{\delta T}{T_0} = & - \int_{\Omega} \frac{1}{T_0} \int_0^T \frac{2}{c_0^2(\mathbf{x}')} \partial_t^2 s(\mathbf{x}', t) \\ & \times \frac{\delta c(\mathbf{x}')}{c(\mathbf{x}')} \bar{s}^\dagger(\mathbf{x}', \mathbf{x}_r, T-t) dt d\mathbf{x}'. \end{aligned} \quad (32)$$

From a comparison of (32) with (14), we infer

$$K(\mathbf{x}, \mathbf{x}_r) = -\frac{2}{T_0 c_0^2(\mathbf{x})} \int_0^T \bar{s}^\dagger(\mathbf{x}, \mathbf{x}_r, T-t) \partial_t^2 s(\mathbf{x}, t) dt. \quad (33)$$

On the basis of eq. (31) and of the definition of Green's function,  $\bar{s}^\dagger$  coincides with the wavefield originated on the membrane by a source

$$\tilde{f}^\dagger(\mathbf{x}, t) = \frac{1}{N} w(T-t) \partial_t s(\mathbf{x}_r, T-t) \delta(\mathbf{x} - \mathbf{x}_r). \quad (34)$$

In analogy with Tromp *et al.* (2005), we call  $\bar{s}^\dagger$  and  $\tilde{f}^\dagger$  ‘adjoint field’ and ‘adjoint source’, respectively. Note that the adjoint source is by definition located at the receiver  $\mathbf{x}_r$ , and that it contains the time-reversed velocity seismogram from the forward synthetic wavefield.

The practical relevance of eqs (31)–(34) becomes apparent when one realizes that (31) can be implemented numerically, feeding  $\tilde{f}^\dagger$  as defined by (34) to a numerical algorithm like our finite-difference membrane scheme, reversed in time. One forward-propagating simulation must be conducted previously, so that  $s(\mathbf{x}_r, t)$  be known (its first and second derivatives with respect to time can be determined numerically): the adjoint source is then entirely defined, and one more run of the finite-difference algorithm is sufficient to determine  $K(\mathbf{x}, \mathbf{x}_r)$ , for all values of  $\mathbf{x}$ . Whatever the source-station geometry, and the complexity of the starting model  $c_0(\mathbf{x})$ ,  $K(\mathbf{x}, \mathbf{x}_r)$  is known after two simulations only.

## 3.3 Some practical considerations

### 3.3.1 Discretization of the adjoint source

Like the initial displacement (5), (6), the adjoint source must be discretized on our membrane grid; let us introduce a discretized version  $\tilde{f}_m^\dagger$  (subscript  $m$  for membrane) of  $\tilde{f}^\dagger$  as defined by eq. (34),

$$\tilde{f}_m^\dagger(\mathbf{x}, t) = \frac{1}{N} w(T-t) \partial_t s(\mathbf{x}_r, T-t) h(\mathbf{x}, \mathbf{x}_r), \quad (35)$$

with

$$h(\mathbf{x}, \mathbf{x}_r) = \begin{cases} 1 & \text{when } \mathbf{x} \text{ is in the same gridcell as } \mathbf{x}_r \\ 0 & \text{elsewhere.} \end{cases} \quad (36)$$

Let us then denote  $\bar{s}_m^\dagger$  the adjoint field generated by  $\tilde{f}_m^\dagger$  on the discretized membrane. Then, by definition of Green's function,

$$\begin{aligned} \bar{s}_m^\dagger(\mathbf{x}, \mathbf{x}_r, T-t) = & \int_0^{T-t} \int_{\Omega} G(\mathbf{x}, \tilde{\mathbf{x}}, T-t-t') \\ & \times \tilde{f}_m^\dagger(\tilde{\mathbf{x}}, t') d\tilde{\mathbf{x}} dt' \end{aligned} \quad (37)$$

$$\begin{aligned} = & \frac{1}{N} \int_0^{T-t} \int_{\text{cell-}\mathbf{x}_r} G(\mathbf{x}, \tilde{\mathbf{x}}, T-t-t') \\ & \times w(T-t') \partial_t s(\mathbf{x}_r, T-t') d\tilde{\mathbf{x}} dt' \end{aligned} \quad (38)$$

$$\begin{aligned} = & \frac{1}{N} \int_0^{T-t} G(\mathbf{x}, \mathbf{x}_r, T-t-t') w(T-t') \\ & \times \partial_t s(\mathbf{x}_r, T-t') dt' \int_{\text{cell-}\mathbf{x}_r} d\tilde{\mathbf{x}} \end{aligned} \quad (39)$$

$$= \bar{s}^\dagger(\mathbf{x}, \mathbf{x}_r, T-t) A_r, \quad (40)$$

with cell- $\mathbf{x}_r$  denoting the surface of the gridcell containing  $\mathbf{x}_r$ , and  $A_r$  its area.

As in practical implementations we shall obtain the ‘discretized’ adjoint wavefield  $\bar{s}_m^\dagger$ , it makes sense to use (40) and replace  $\bar{s}^\dagger$  with  $\bar{s}_m^\dagger/A_r$  in (33), to find an expression for the discrete version  $K_m$  of sensitivity kernels

$$K_m(\mathbf{x}, \mathbf{x}_r) = -\frac{2}{A_r T_0 c_0^2(\mathbf{x})} \int_0^T \bar{s}_m^\dagger(\mathbf{x}, \mathbf{x}_r, T-t) \partial_t^2 s(\mathbf{x}, t) dt. \quad (41)$$

### 3.3.2 Waveform filtering

While we are interested in the propagation of one mode (one frequency of the dispersive surface wave packet) at a time, our membrane analogue is excited over a range of frequencies, depending on the initial conditions (5) and (6). To isolate the mode of interest, we bandpass-filter the solution using as centre frequency the frequency of the mode of interest, and a half-bandwidth of 2.5 mHz. Spetzler *et al.* (2002) average their sensitivity kernels over the same bandwidth, to account for the fact that single-frequency phase-velocity measurements are not possible, owing to the finite sampling of seismograms and to the finite parametrization of the dispersion curve in the measurement process. The selected value for bandwidth also coincides with the spacing between splines parametrizing the measured dispersion curves in Ekström *et al.* (1997, fig. 1), which could be taken as a rough estimate of the accuracy of said dispersion curves (Boschi 2006).

As noted in Section 2.1, each membrane simulation provides the surface wave potential  $s$  associated with one specific surface wave mode. Strictly speaking, only the component of  $s$  at the frequency of the mode of interest is then physically meaningful. The relatively large bandwidth of our bandpass filter leads, however, to an effect similar to that found analytically by Spetzler *et al.* (2002), reducing the amplitude of kernels’ sidelobes. Filtering can be equivalently applied to the scalar potential  $s$  found by numerical integration on the membrane, or to the initial conditions (5) and (6). We have experimented with both approaches, obtaining practically coincident results. The latter option is naturally more efficient.



## 4 RESULTS

The discussion that follows is limited to phase-velocity kernels of intermediate-to-long period Love waves; we expect that a similar procedure is valid, and the same qualitative results hold for the case of Rayleigh waves.

### 4.1 Sensitivity kernels for a homogeneous (spherical-Earth) starting model

#### 4.1.1 Comparing analytically and numerically determined sensitivity kernels

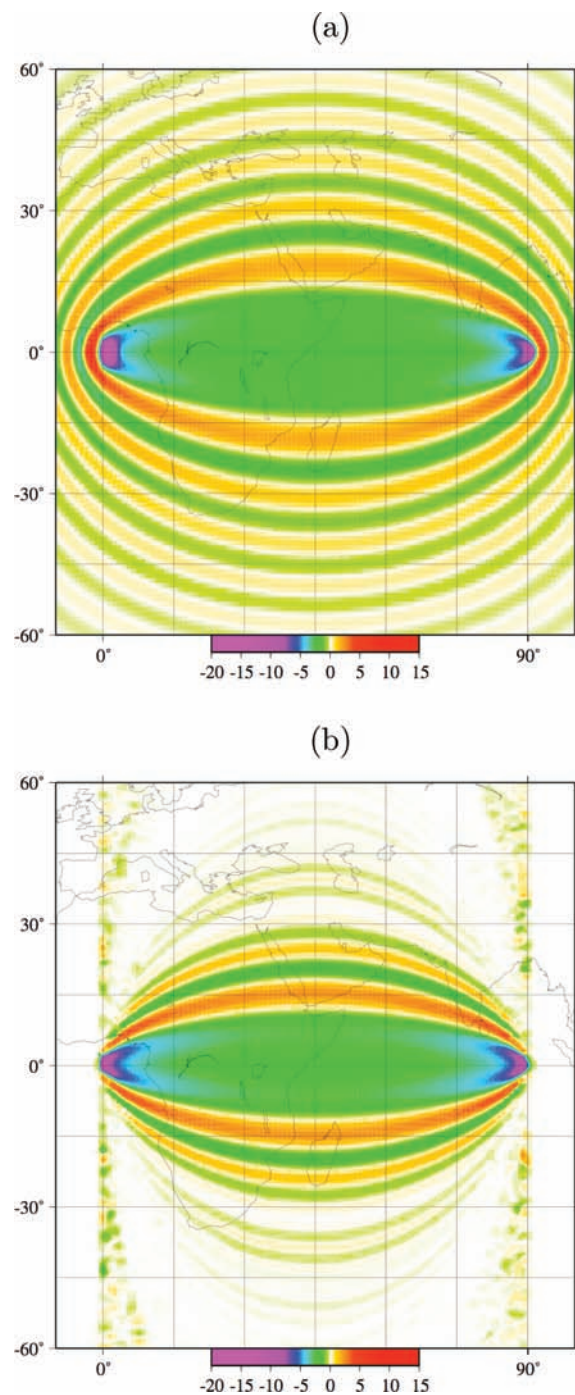
In Fig. 5, we compare a sensitivity kernel calculated numerically in the membrane approach, with one calculated analytically from Boschi's (2006) implementation of Spetzler *et al.*'s (2002) eq. (16), based on Snieder & Nolet's (1987) single-scattering approach. Source ( $0^\circ\text{N}$ ,  $0^\circ\text{E}$ ) and receiver ( $0^\circ\text{N}$ ,  $90^\circ\text{E}$ ) locations, and surface wave period (150 s) are the same in both cases. The background value for phase velocity coincides with the Love-wave, PREM-based (Dziewonski & Anderson 1981) value. Numerical kernel values are calculated on the cell midpoints (as described in Section 2.3) of our grid and then interpolated for plotting with  $1^\circ$  spacing in both latitude and longitude. Analytical values are likewise averaged over each cell and interpolated for plotting.

As pointed out by Favier *et al.* (2004) and Boschi (2006), Spetzler *et al.*'s (2002) and other Born-theory formulations (e.g. Zhou *et al.* 2004; Yoshizawa & Kennett 2005) involve a far-field approximation of the unperturbed solution; sensitivity kernels found in this approach are necessarily singular at source and receiver. This explains the unphysical behaviour of analytical kernels evident from Fig. 5(b) at longitudes around  $0^\circ$  and  $90^\circ$ . Furthermore, analytical kernels in Fig. 5 are zero at longitudes larger than the epicentral distance, and at any location at a negative azimuth from the source; this is an effect of simplifications in Spetzler *et al.*'s (2002) procedure—a fictitious feature that we do not find in expressions for analytical kernels later derived by Zhou *et al.* (2004) and Yoshizawa & Kennett (2005).

#### 4.1.2 Comparing sensitivity kernels found via the 'direct' versus adjoint approach

In Fig. 6, a comparison between cross-sections of numerical kernels for Love waves at 150 s period is shown. Numerical kernel values are either calculated with the 'direct' approach or the adjoint method, both described in Section 3. We employ in both cases the same numerical mesh and homogeneous background model. For the 'direct' approach, we apply a relative perturbation  $\gamma$  of  $-0.2$  per cent. Source and station are at a fixed epicentral distance and placed as in Fig. 5. Cross-sections of both kernels are taken at distinct longitudes. The 'direct' calculation takes about 6 min on 16 processors to produce a complete numerical-direct kernel over the whole sphere (using the reduction from Section 3.1.1). With the adjoint method, the complete numerical-adjoint kernel is obtained in about 2 min on a single processor. As can be seen, the two kernels are practically identical. This is an important demonstration of the internal consistency of our approaches.

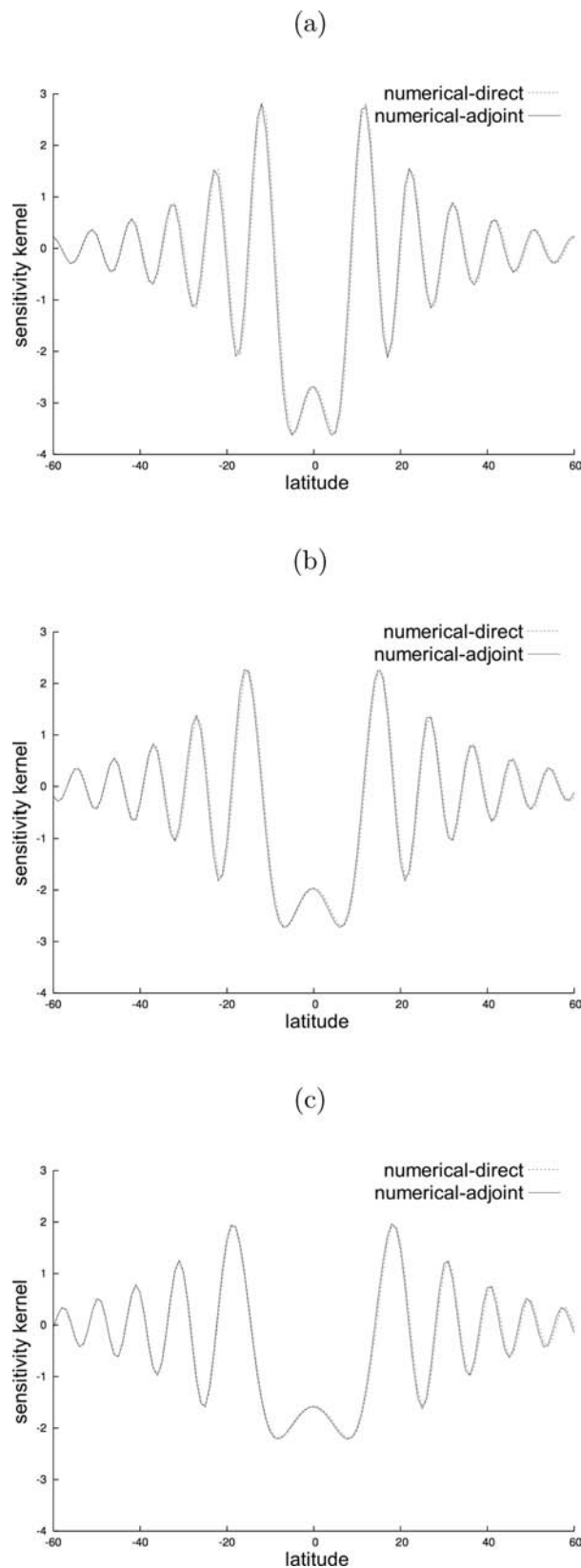
In Fig. 7, we explore how membrane sensitivity kernels vary as a function of source-station distance; the area of all Fresnel zones expands as already described by, for example, Spetzler *et al.* (2002, fig. 2b). For stations located closer to the epicentre, overall higher sensitivity is found. Fig. 8 shows the dependence of sensitivity ker-



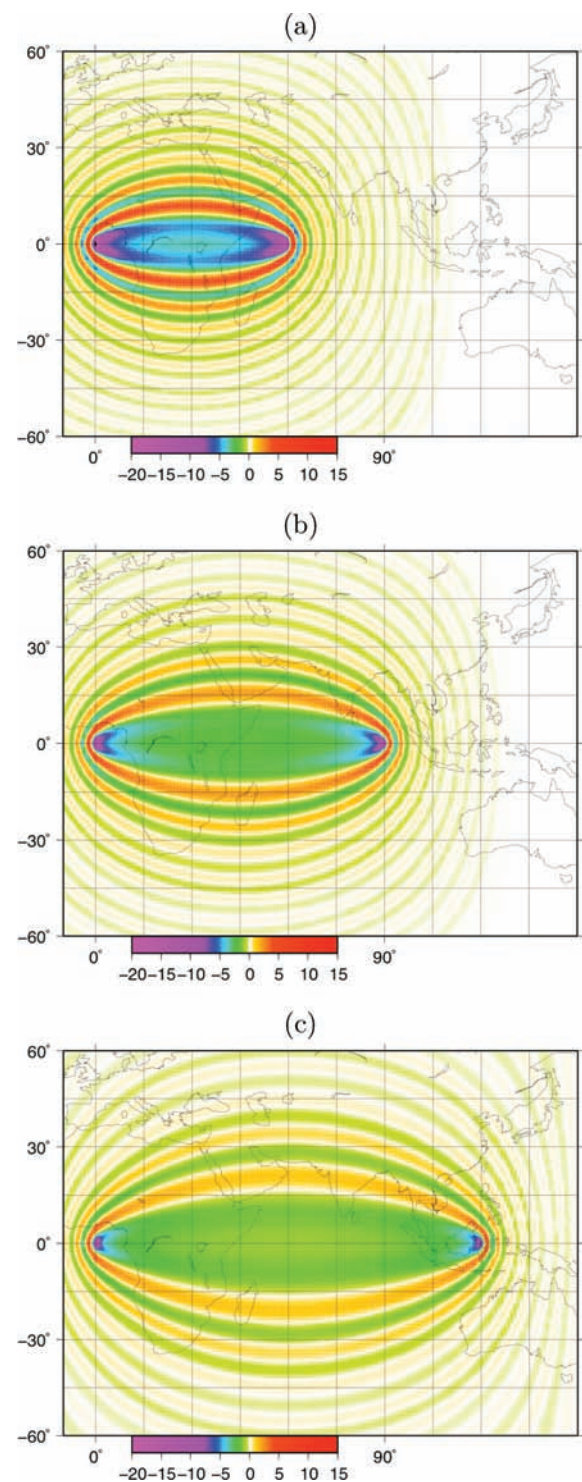
**Figure 5.** Sensitivity kernel for Love waves at 150 s, spherical earth model (homogeneous phase velocity), source at ( $0^\circ\text{N}$ ,  $0^\circ\text{E}$ ) and receiver at ( $0^\circ\text{N}$ ,  $90^\circ\text{E}$ ), (a) calculated numerically in the 'direct' approach; (b) calculated implementing the analytical formula of Spetzler *et al.* (2002).

nels on wave period at a fixed epicentral distance. The central lobes of the kernels increase with increasing period. This is qualitatively confirmed by analytical results (e.g. Spetzler *et al.* 2002, fig. 2a).

Due to the memory-intensive storage of the numerical grid, our current computer hardware prevents us from running simulations with periods  $< 75$  s. The numerical scheme we apply requires about 10 nodes per wavelength for an accurate representation of the wave phenomena on the membrane (see Section 2.3); thus, for shorter periods we would need a grid spacing with less than 17 km distance



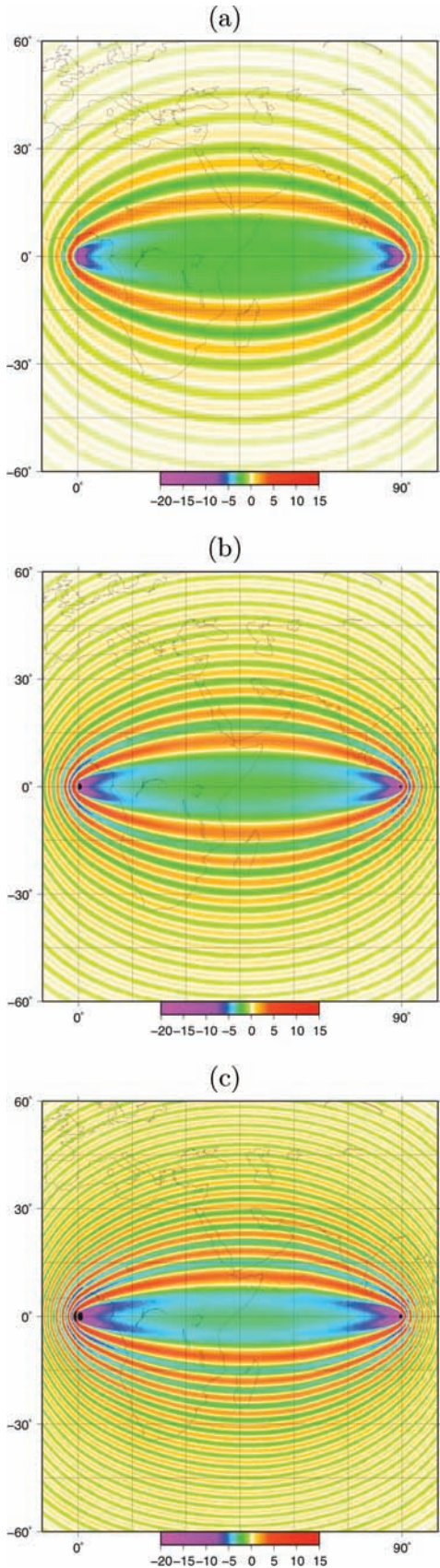
**Figure 6.** Cross-sections of sensitivity kernels for Love waves at 150 s calculated via the ‘direct’ approach (dashed lines) and adjoint method (solid lines), with source at ( $0^\circ\text{N}$ ,  $0^\circ\text{E}$ ) and receiver at ( $0^\circ\text{N}$ ,  $90^\circ\text{E}$ ). Sensitivity values are plotted as a function of latitude only, along three chosen meridians, namely (a)  $10^\circ$ , (b)  $20^\circ$  and (c)  $45^\circ$ .



**Figure 7.** Numerical-adjoint kernels derived for 150 s Love waves in a homogeneous starting model. The source is located at ( $0^\circ\text{N}$ ,  $0^\circ\text{E}$ ), the receiver at  $0^\circ\text{N}$  and (a)  $60^\circ\text{E}$ , (b)  $90^\circ\text{E}$  and (c)  $120^\circ\text{E}$ .

which leads to more than two million grid cells. With 2 GB RAM memory per cluster node we are bound by this limitation. A work-around could consist of accessing the mesh points through I/O with a corresponding file storage. This would considerably slow down the computation process.





**Figure 8.** Numerical-adjoint kernels for Love waves at (a) 150 s, (b) 100 s and (c) 75 s periods, in a homogeneous starting model. The source is located at (0°N, 0°E), the receiver at (0°N, 90°E).

#### 4.2 Sensitivity kernels for a laterally heterogeneous starting model

Our method naturally allows us to calculate phase-velocity sensitivity kernels associated with a laterally heterogeneous starting model (background Earth). Computing a single numerical kernel in the adjoint approach takes exactly the same time (2 min on one of our processors) regardless whether the background Earth is homogeneous or heterogeneous. In the limit of our smooth-Earth assumption, we explore the impact of background heterogeneities on the properties of surface wave kernels.

Fig. 9 shows a starting model for 150 s Love wave phase velocity. We derived it based on local normal-mode theory (Boschi & Ekström 2002), from an earth model consisting of the 3-D crustal model Crust-2.0 (Bassin *et al.* 2000), overlying a 1-D, radially isotropic profile of the mantle as in Boschi *et al.* (2004). The map in Fig. 9 represents a rough guess of crustal effects on long-period Love wave propagation, and is independent of surface wave observations like those we will invert in the following. In order to expect significant differences, it has been filtered to harmonic degrees  $\leq 40$  (thus allowing spatial wavelengths close to the ones of Love waves at 150 s). Our implementation is only physically meaningful in the smooth-Earth approximation (Section 2.1); the algorithm remains stable in the presence of strong gradients, but the formulation of the membrane approach depends upon lateral smoothness.

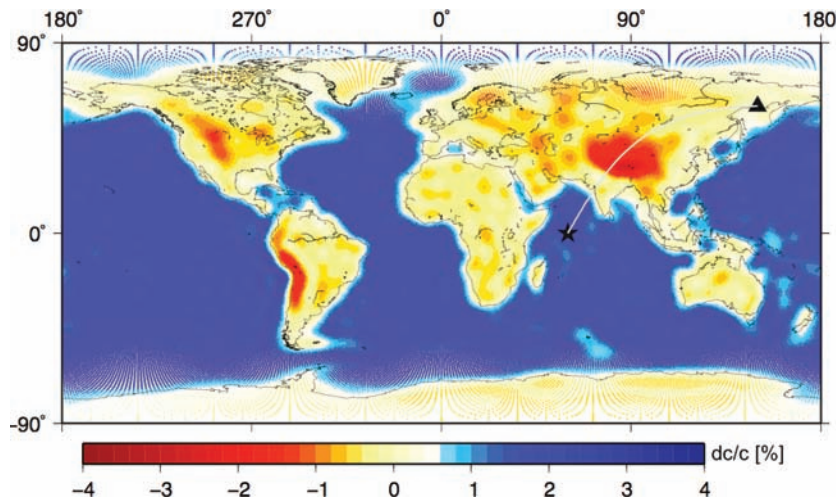
We show in Fig. 10(a) a 150 s Love wave kernel based on the model of Fig. 9, and a source-station geometry that should maximize the effect of the starting model's strongest phase-velocity heterogeneity, located in the Himalaya region with approximately  $-3$  per cent relative phase-velocity perturbation. The homogeneous-Earth kernel associated with the same source and receiver is shown in Fig. 10(b), and the difference between the two in Fig. 10(c). The two kernels coincide in the first Fresnel-zone (main lobe), while significant differences are apparent in the sidelobes. Strongest differences are found to the southeast of the great circle path, corresponding to higher gradients in phase velocity (transition from a continental to an oceanic region).

#### 4.3 A test of the first-order scattering approximation

Born theory is a single-scattering theory, that is, it neglects the interaction of scattered wavefields with other heterogeneities. In practice, the linearized, Born-theoretical eq. (14) implies that the effect of multiple heterogeneities be equal to the sum of the independently calculated effects of each heterogeneity. In forward calculations made on a smooth earth model, our numerical algorithm naturally accounts for multiple scattering. Hence, while we do not attempt to formulate an inverse problem accounting for multiple scattering, we can perform a set of forward calculations to evaluate its relevance, and the associated inaccuracy of the linearized Born approximation.

Let us denote  $\delta\Phi_{jk}$ , the phase anomaly at the receiver due to the presence of two scatterers centred at  $(\vartheta_j, \varphi_j)$  and  $(\vartheta_k, \varphi_k)$  (i.e. extending over the whole cell of our hexagonal grid, centred at  $\vartheta_j, \varphi_j$  or  $\vartheta_k, \varphi_k$ ), in an otherwise homogeneous background model. We compute this quantity in two ways: (i) by cross-correlation between the *numerical* solution found propagating membrane waves in the presence of two scatterers, and the one found on a homogeneous membrane; (ii) making use of previously calculated kernels  $K(\vartheta, \varphi)$  to implement

$$\delta\Phi_{jk} = K(\vartheta_j, \varphi_j)\Phi\gamma A_j + K(\vartheta_k, \varphi_k)\Phi\gamma A_k, \quad (42)$$



**Figure 9.** 150 s Love-wave phase-velocity map based on Crust-2.0 (Bassin *et al.* 2000) and an isotropic upper-mantle model. The phase-velocity heterogeneities are represented by a spherical harmonic expansion with degrees  $l \leq 40$ , resulting in a low-pass filtered phase-velocity map (the membrane analogue has physical meaning in a smooth Earth regime). Phase-velocity values are projected onto our hexagonal grid, and plotted as coloured dots at the centre of each of its cells (122 882 total cells). Source and receiver are denoted by the star and triangle connected by the source–receiver great circle, which crosses the Tibet anomaly. Phase-velocity anomalies are given in percent with respect to PREM.

based upon eq. (14), where  $\Phi$  is the homogeneous reference phase,  $A_j$ ,  $A_k$  are the areas of the cells to which  $(\vartheta_j, \varphi_j)$  and  $(\vartheta_k, \varphi_k)$ , respectively, belong,  $\gamma$  is the phase-velocity perturbation, which for simplicity we choose to have the same value at both scatterers.

In the approach (i), multiple scattering is implicitly accounted for; in the approach (ii) it is neglected. The linearized Born approximation would suggest that both values of  $\delta\Phi_{jk}$  are identical; differences between the resulting phase anomalies are the effect of multiple scattering. We perform a number of simulations with: source and receiver located at  $(0^\circ, 0^\circ)$  and  $(0^\circ, 90^\circ)$ , respectively; one scatterer located at  $45^\circ$  longitude, and at latitudes varying between  $-60^\circ$  and  $60^\circ$ ; a second scatterer at the same latitude, and at a longitudinal distance of  $+3^\circ$ ,  $+6^\circ$  or  $+9^\circ$  to the first. For each considered couple of scatterers, we find phase anomaly both by ‘direct’ numerical calculation [approach (i)] and by eq. (42) [approach (ii)]; we plot in Fig. 11 the difference between the resulting values of  $\delta\Phi$  as a function of scatterer-latitude.

As a general rule, we find that the effects of multiple scattering on seismic phase are small and the linearization in eq. (14) is valid. The largest discrepancy in  $\delta\Phi$  amounts to  $\sim 0.15$  per cent of the maximum  $\delta\Phi$  predicted by eq. (42) and corresponds to the smallest distance ( $3^\circ$ ) between the two scatterers. These values are so small that the shape of the curves is significantly affected by the irregularity of the grid: gridpoints do not precisely align along parallels, but can be shifted up to  $\pm 0.5^\circ$  in our mesh and the curves in Fig. 11 are jagged as a result. Multiple scattering becomes even less relevant for larger interscatterer distances ( $6^\circ$  and  $9^\circ$  in our experiment). The latter result is to be expected, as the energy of the scattered wavefield at a given point decreases with increasing distance from the scatterer, by simple geometrical spreading. We infer that in a smooth-Earth regime the Born linearization is valid in the tomographic determination of phase-velocity anomalies. It would become less reliable at longer times/higher orbits.

#### 4.4 Application to fundamental-mode surface wave tomography

The inverse problem of identifying a phase-velocity map from dispersion measurements compiled from a large set of source–receiver

pairs is typically reduced to solving a linear system

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{d} \quad (43)$$

(e.g. Boschi 2001), where the entries of  $\mathbf{x}$  are the unknown coefficients of the phase-velocity map to be determined and those of  $\mathbf{d}$  the measured relative phase anomalies  $\delta\Phi/\Phi$ . As a result of eq. (14), and employing a pixel parametrization (Boschi 2006),

$$A_{ij} = \int_{j^{\text{th}} \text{ pixel}} K_i^{\text{th datum}}(\omega, \vartheta, \varphi) d\Omega. \quad (44)$$

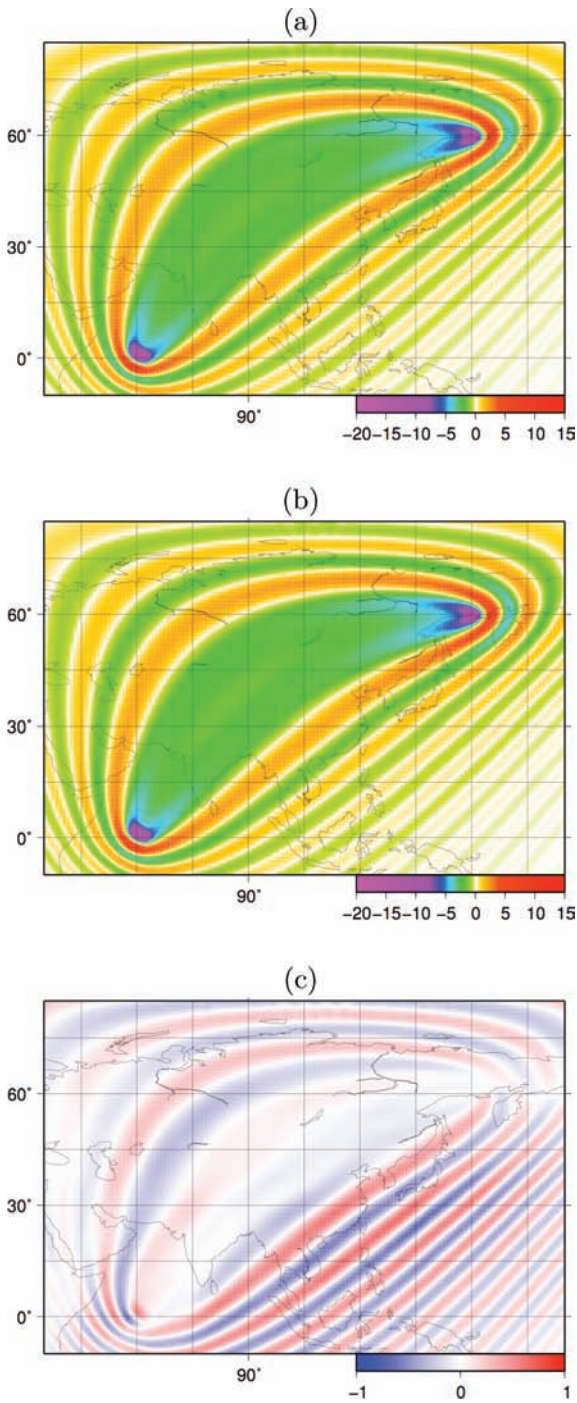
We invert the database of Ekström *et al.* (1997), updated as described by Boschi & Ekström (2002), implementing (44) and least-squares solving (43) with the algorithm of Boschi (2006), but calculating  $K(\omega, \vartheta, \varphi)$  in different ways and different starting models, as described below. Our phase-velocity maps are linear combinations of equal-area pixel functions as in Boschi *et al.* (2006). The coverage (see Fig. 12 for 150 s Love-wave observations) and resolving power of the Harvard dispersion database have been evaluated in earlier publications (e.g. Ekström *et al.* 1997; Carannante & Boschi 2005).

##### 4.4.1 Tomography with a homogeneous starting model

We follow the procedure described in Sections 3.2 and 3.3 to calculate sensitivity kernels  $K(\omega, \vartheta, \varphi)$  defined by eq. (14) [strictly speaking, what we find and use is their ‘discretized version’  $K_m(\omega, \vartheta, \varphi)$  defined by (41)]. So long as the starting model is homogeneous, the function  $K(\omega, \vartheta, \varphi)$  changes if the source–receiver distance changes, but is not affected by changes in the locations of source and receiver: following Boschi (2006), we find  $K(\omega, \vartheta, \varphi)$  at a discrete set of epicentral distances ranging from  $20^\circ$  to  $179^\circ$ , with  $1^\circ$  increments. We later spline-interpolate (Press *et al.* 1992) calculated  $K(\omega, \vartheta, \varphi)$ ’s to find  $K(\omega, \vartheta, \varphi)$  for any epicentral distance (Boschi 2006). In analogy with Spetzler *et al.* (2002) or Boschi (2006), we neglect source-mechanism variations for different events in the database, and use the source term (5), (6).

After implementing eq. (44) for the entire database at a chosen surface wave mode, we least-squares invert (43) a number of times, varying the value of the roughness-damping parameter (Boschi





**Figure 10.** Sensitivity kernels derived with the adjoint method for 150 s Love waves in (a) homogeneous and (b) heterogeneous (Fig. 9) starting phase-velocity models. (c) Difference between (a) and (b).

2006; Boschi *et al.* 2006); no other regularization constraint is applied. The resulting L-curve, or plot of misfit versus normalized roughness of the solution, as defined by Boschi (2006), is shown in Fig. 13(a). While we experimented with a variety of surface wave modes and tomographic parametrization, we shall limit our discussion to 150 s Love wave data inverted on a grid of equal area pixels with surface extent  $3^\circ \times 3^\circ$  at the equator.

We repeat this exercise, employing alternatively ray-theoretical sensitivity kernels, and analytical Born-theoretical ones based on

Spetzler *et al.* (2002). We find that the L-curves resulting from the three approaches are qualitatively similar, as also noted by Boschi (2006). We identify from Fig. 13(b) solutions corresponding to equal curvature of the associated L-curves, and roughness comparable to that of, for example, Ekström *et al.* (1997). Associated phase-velocity maps are shown in Fig. 14. The results of the three approaches are remarkably similar; the only discrepancies worth mentioning are perhaps two fast anomalies of small lateral extent, in the southeastern part of the Chinese Gansu province and in the Andes, present in the analytical-Born-theory solution, but not in the other two.

The most time-consuming part of our experiment was the calculation of sensitivity kernels at the mentioned set of 160 epicentral distances, for one surface wave mode (150 s Love waves). Applying the adjoint method on one processor, this would last about 5 hr. Computing  $A_{ij}$  for 16 624 observations of  $\delta \Phi/\Phi$  takes another 5 hr. Finally, least-squares solving the resulting linear inverse problem by means of LSQR, with  $\mathbf{A}$  relatively dense, for a large set of roughness-damping parameter values takes a further 3 hr.

#### 4.4.2 Tomography with laterally heterogeneous starting models

As discussed in Section 4.2, in a laterally heterogeneous Earth the form of a sensitivity function  $K(\omega, \vartheta, \varphi)$  depends not only on epicentral distance, but also on the specific locations of source and receiver. For this reason,  $K(\omega, \vartheta, \varphi)$  has to be calculated for each source–receiver pair, that is, for each observation in the database. At the speed of currently available hardware, this would be practically impossible if the ‘direct’ approach algorithm of Section 3.1 were used. It becomes feasible when the procedure described in Section 3.2 is applied, thanks to the gain in speed achieved with the adjoint approach.

We limit ourselves, again, to Love-wave phase-anomaly observations at 150 s, and use as a starting phase-velocity model the one of Fig. 9, based on the crustal model Crust-2.0 and an isotropic upper-mantle model (Section 4.2). We next

(i) multiply the matrix  $\mathbf{A}$  found in the homogeneous-model case by the vector of starting-model coefficients, and subtract the resulting phase-anomaly vector from the data; let us denote  $\mathbf{d}'$  the resulting, corrected phase-anomaly vector;

(ii) following the method described in Section 4.2, compute a sensitivity kernel for each of the 16 624 summary observations (150 s Love waves only) available in the Harvard database; let us denote  $K'_{i^{\text{th}} \text{ datum}}(\omega, \vartheta, \varphi)$  the kernel associated with the  $i$ -th observation;

(iii) after a kernel is calculated, augment the matrix  $\mathbf{A}'$  accordingly; the  $ij$  entry of  $\mathbf{A}'$  is naturally defined

$$A'_{ij} = \int_{j^{\text{th}} \text{ pixel}} K'_{i^{\text{th}} \text{ datum}}(\omega, \vartheta, \varphi) d\Omega, \quad (45)$$

with the same pixel grid used in previous inversions (approximately equal-area pixels,  $3^\circ \times 3^\circ$  at the equator);

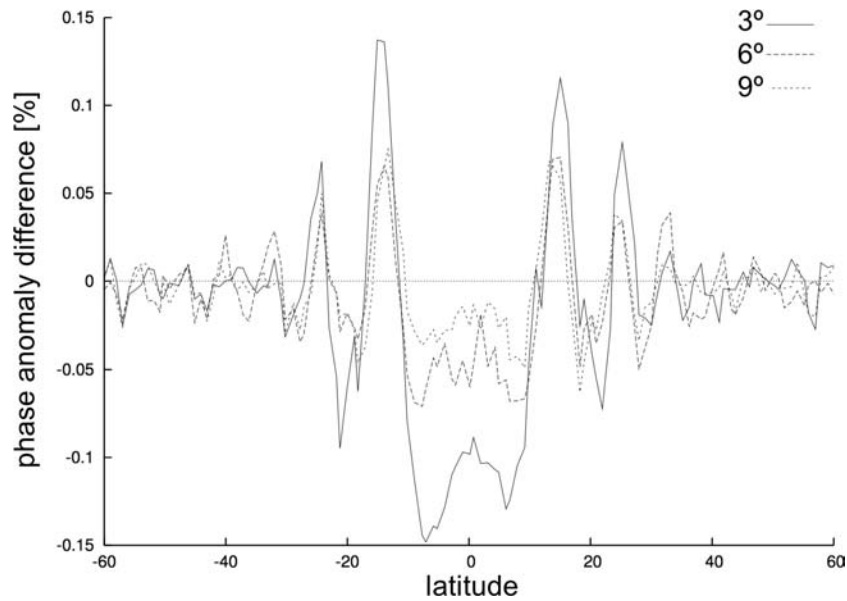
(iv) iterate the two previous steps until  $\mathbf{A}'$  accounts for all available data;

(v) dubbed  $\mathbf{x}'$  the coefficients of perturbations to the starting model of Fig. 9, least-squares solve the linear inverse problem

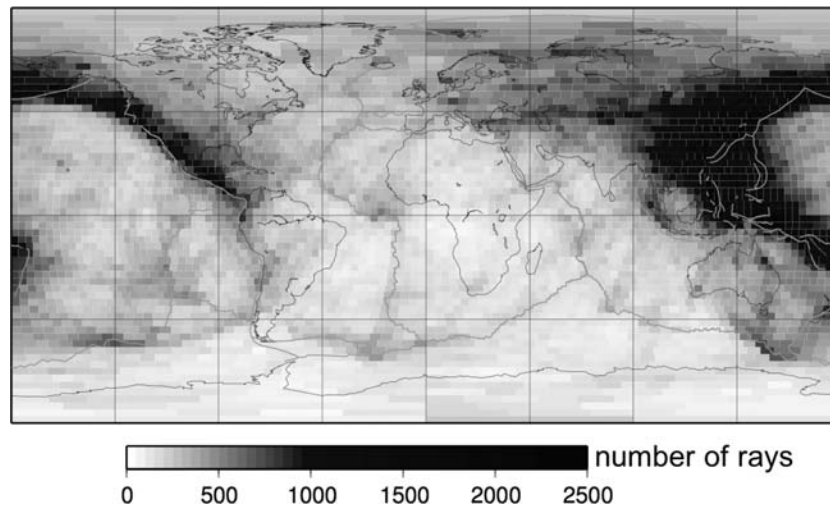
$$\mathbf{A}' \cdot \mathbf{x}' = \mathbf{d}' \quad (46)$$

repeatedly, for a wide range of values of the roughness damping parameter (again, no other regularization constraint is applied);

(vi) conduct a trade-off analysis to identify a preferred solution model, whose regularization is compatible with that applied in earlier experiments, so that the results can be compared (Boschi 2006).



**Figure 11.** Effects of multiple scattering. We use the membrane-wave method to calculate the phase anomaly  $\delta\Phi_i^{\text{num}}$  resulting from two scatterers, located at the same latitude and longitudinal distances  $3^\circ$ ,  $6^\circ$  and  $9^\circ$ , and source and receiver at  $(0^\circ\text{N}, 0^\circ\text{E})$  and  $(0^\circ\text{N}, 90^\circ\text{E})$ , respectively. After each simulation, we subtract  $\delta\Phi_i^{\text{num}}$  from the phase anomaly  $\delta\Phi_i^{\text{Born}}$  by simple Born theory (no multiple scattering, but the sum of the individual effects of each scatterer) and normalize the result to the maximum value of  $\delta\Phi_i^{\text{Born}}$  from all our multiple-scattering simulations. The resulting quantity  $(\delta\Phi_i^{\text{Born}} - \delta\Phi_i^{\text{num}})/\max\{\delta\Phi_i^{\text{Born}}\}$  is plotted in percent vs. scatterer latitude, with a separate curve for each value of longitudinal distance between scatterers ( $3^\circ$ : solid line,  $6^\circ$ : long-dashed line,  $9^\circ$ : short-dashed line).

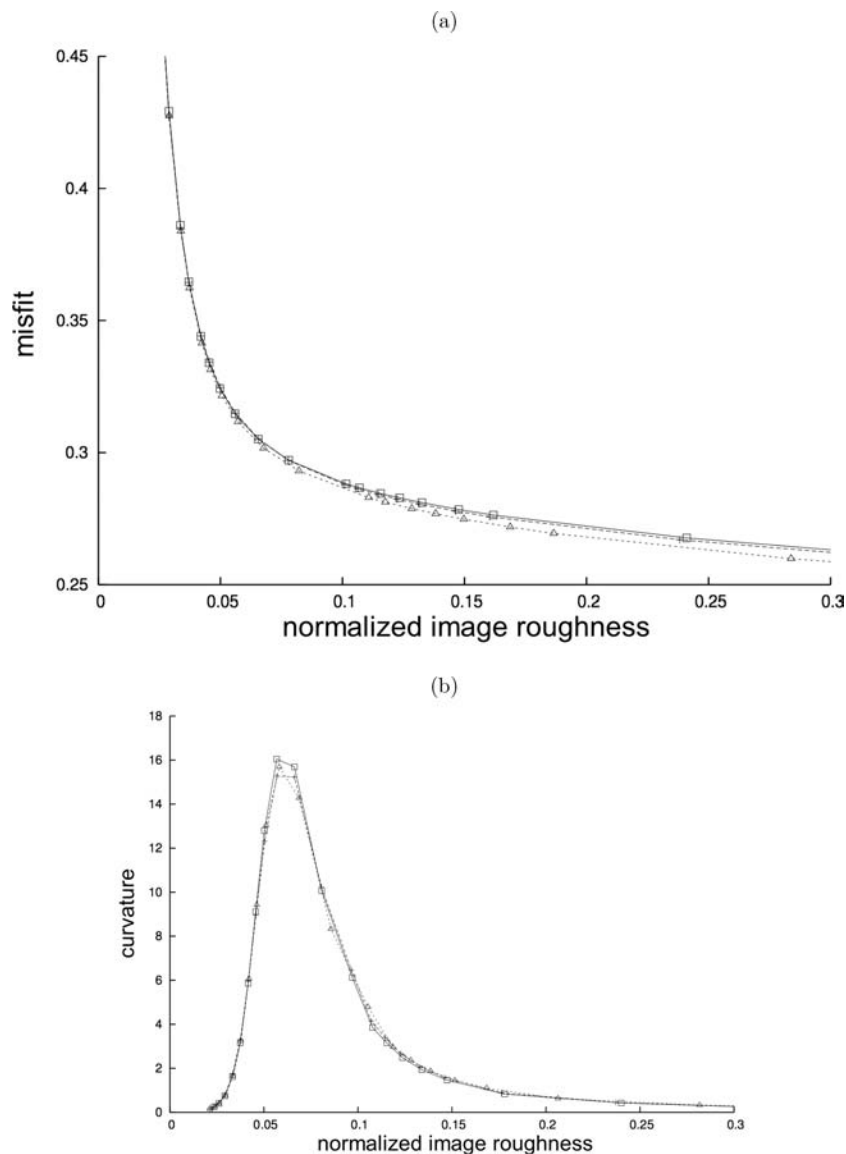


**Figure 12.** Ray-theoretical hitcount map (number of rays crossing each pixel) from the Harvard database, 150 s Love wave observations only. The equal-area pixel parametrization ( $3^\circ \times 3^\circ$  at the equator) is the same used in our inversions.

The L-curve and its curvature resulting from the trade-off analysis (vi) are shown in Fig. 15(a) and (b), respectively, as functions of normalized image roughness. The L-curve of Fig. 13 derived by numerical-adjoint kernels is also shown in Fig. 15 for comparison. Starting the inversion from a heterogeneous model, derived from Crust-2.0 and a spherically symmetric upper-mantle model (Fig. 9), leads to solutions with a worse datafit than the homogeneous-starting-model ones of equal roughness, suggesting that the chosen heterogeneous model does not describe sufficiently well the propagation of Love waves at 150 s period. As in this approach the starting model has non-zero roughness, the corresponding L-curve does not tend to zero for decreasing model complexity, but converges to the roughness value of the initial starting model.

We repeated this exercise, using as a starting model the one shown in Fig. 14(c), derived from the inversions using numerical-adjoint kernels. This represents a first iteration step in a non-linear inversion scheme where the zeroth iteration starts with a homogeneous phase-velocity model. Steps (ii), (iii) and (iv) are the most time-consuming and took about 22 hr using 16 processors in our cluster. Step (v) takes about 3 hr on a single processor, as in the homogeneous-model case: this is not surprising since  $\mathbf{A}$  and  $\mathbf{A}'$  are equally dense. The solutions obtained have a slightly better fit to the data than from the zeroth iteration. We iterated the process one more time, to find only insignificant changes in model and misfit.

The points of equal curvature in Fig. 15(b) (dotted line/crosses, dashed line/circles) identify our preferred solutions, which we show



**Figure 13.** Trade-off analysis for the phase-velocity inversions of Section 4.4.1 (homogeneous starting model). (a) L-curves for solutions derived from ray theory (dotted line, triangles), analytical Born theory (dashed line, pluses), and numerical-adjoint kernels (solid line, squares). (b) Curvature of the curves shown in (a). Image roughness is defined and normalized as in Boschi (2006).

in Fig. 16. We first combine  $x'$  with the starting model (of Fig. 9 and Fig. 14c, respectively), so that heterogeneities are again defined with respect to the value of 150 s Love-wave phase velocity predicted by PREM. Only few, small differences between the maps of Fig. 16 and those of Fig. 14 are apparent. Namely, in the top panel of Fig. 16 a fast, southeastern Atlantic anomaly can now be distinguished from the fast anomalies corresponding to cratons in western and southern Africa; the same happens in South America, where the fast anomaly in the southern Atlantic Ocean is more clearly separated from the Brazilian one. Using Crust-2.0 as a starting crustal model in a 3-D, ray-theoretical inversion, and then computing phase-velocity maps associated with the resulting 3-D shear velocity model, a similar effect is observed (unpublished result by L. Boschi, 2006, based on the method of Boschi & Ekström 2002 and Boschi *et al.* 2004).

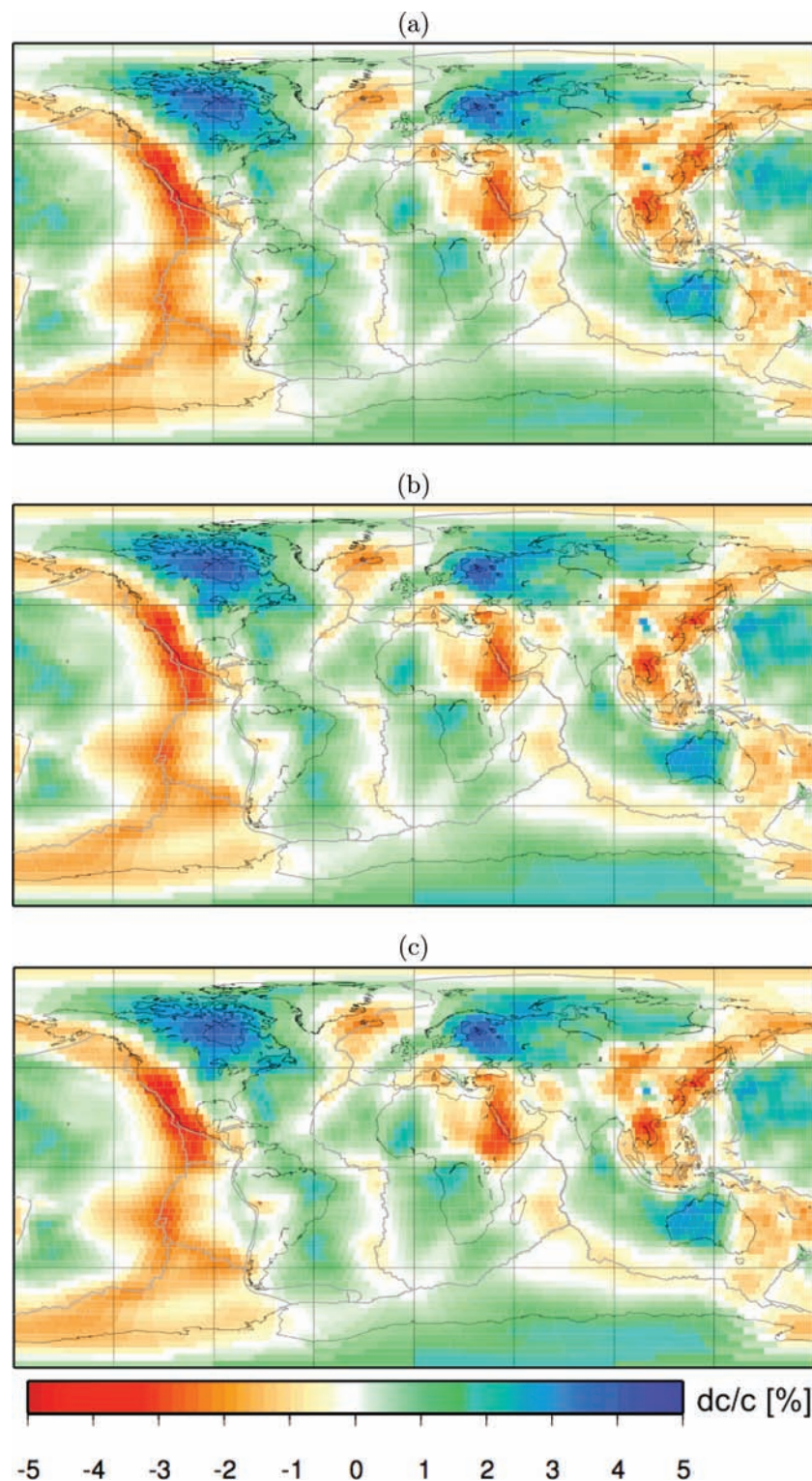
We show in Fig. 17 the power spectrum up to degree 40 of the phase-velocity maps in Figs 14 and 16. All spectra strongly resemble each other, particularly at lower harmonic degrees (longer spatial wavelengths). Note that the phase-velocity maps obtained

from laterally heterogeneous starting models (based on Crust-2.0 and an isotropic upper-mantle model; and based on an initial inversion from PREM) show slightly higher spectral values at degrees  $>5$ .

## 5 CONCLUSIONS

We model surface wave propagation in a smoothly heterogeneous Earth, implementing the wave equation numerically on a spherical membrane (zero thickness). The numerical method utilizes a finite-difference scheme specifically designed for the spherical grids. In comparison with existing techniques, our approach has both advantages and disadvantages. It is less accurate than the fully 3-D, numerical solution of the Earth's equations of motion (e.g. Komatitsch *et al.* 2002; Capdeville *et al.* 2003) in that it requires the smooth-Earth/no-mode-coupling approximation to be made, but an order of magnitude faster. Unlike the analytical approaches like those of, for example, Spetzler *et al.* (2002); Zhou *et al.* (2004) and Yoshizawa





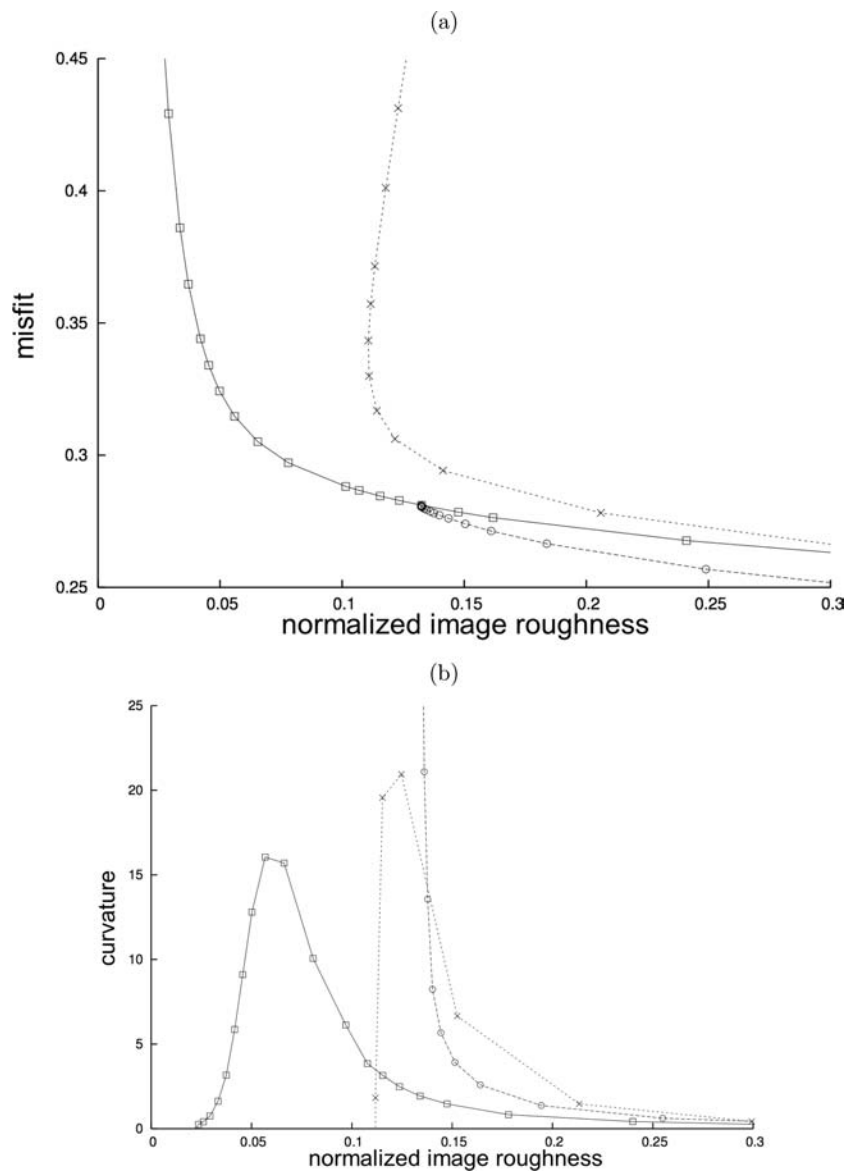
**Figure 14.** Phase-velocity maps from the inversions of section 4.4.1 (homogeneous starting model). We compare solutions found from (a) ray theory, (b) analytical Born theory, (c) numerical-adjoint kernels. Phase anomalies are in percent with respect to the value predicted by PREM.

& Kennett (2005), it does not involve any far-field approximation (Favier *et al.* 2004), and accounts for some of the non-linearities of wave propagation in a realistic medium (Tanimoto 1990). Zhou *et al.*'s (2004) method, on the other hand, considers also mode coupling. Both Zhou *et al.*'s (2004) and Yoshizawa & Kennett's (2005)

methods account for the effects of seismic source radiation, which our membrane analogue does not.

The high speed of our finite-difference membrane wave algorithm, combined with the application of the adjoint method (e.g. Tarantola 1984; Tromp *et al.* 2005), allowed us to use our

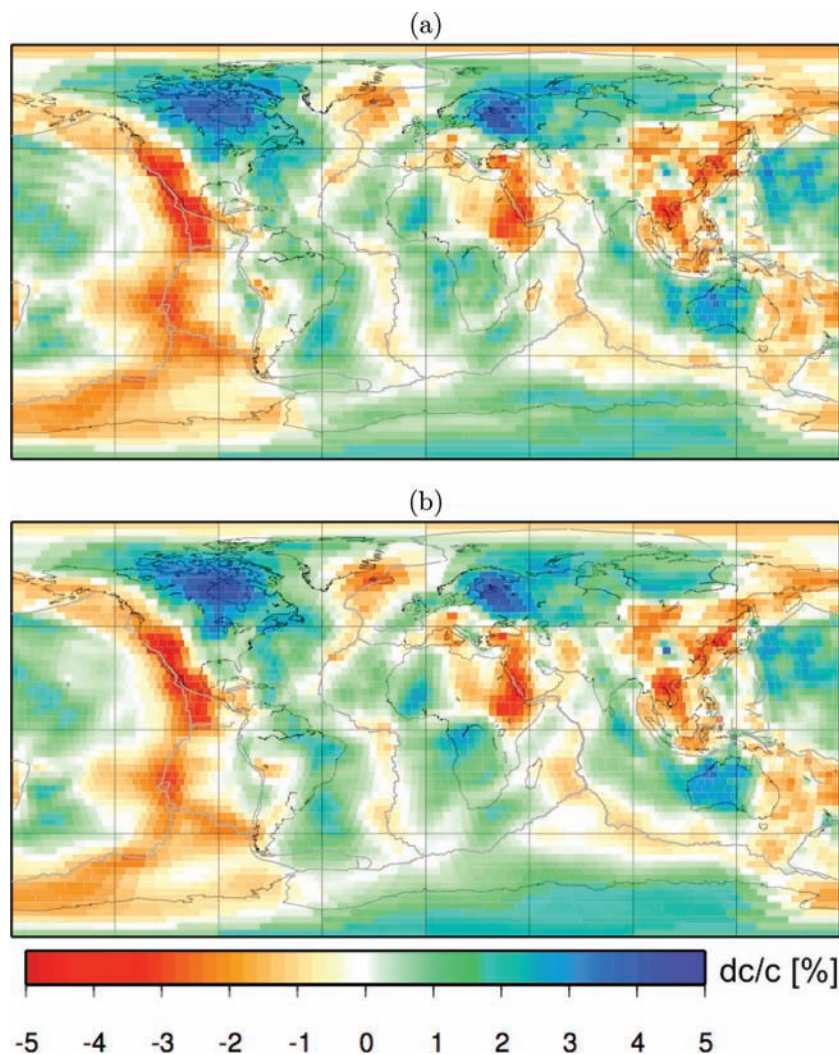




**Figure 15.** Trade-off analysis for the phase-velocity inversions of Sections 4.4.1 and 4.4.2 (heterogeneous starting model). (a) L-curves for solutions derived using numerical-adjoint kernels based on three different starting models: a homogeneous starting model (solid line, squares) (curve repeated from Fig. 13); the heterogeneous starting model shown in Fig. 9, which itself was derived from Crust-2.0 and an isotropic upper-mantle model (dotted line, crosses); the heterogeneous starting model shown in Fig. 14c, which itself was derived using numerical-adjoint kernels (dashed line, circles). (b) Curvature of the curves shown in (a).

membrane analogue to derive sensitivity kernels relating surface wave phase-anomaly data to phase-velocity heterogeneities. We computed sensitivity kernels using two different approaches—one employing adjoint methods, the other using a large set of direct calculations (no backpropagation)—and found coincident results (Fig. 6). We then calculated sensitivity kernels both in a homogeneous and a laterally heterogeneous starting model of phase velocity, and employed them in a set of global inversions of the Harvard dispersion database (e.g. Ekström *et al.* 1997). Kernels calculated in this way are free from the far-field approximation often used in analytical Born theory (e.g. Dahlen *et al.* 2000; Spetzler *et al.* 2002; Boschi *et al.* 2006). Fundamental-mode tomographic images and trade-off analyses (Figs 13 and 14) derived from different approaches (ray theory, analytical Born theory, numerical Born kernels) are approximately coincident (see also Fig. 17).

As explained in Section 2, our membrane wave formulation of surface wave propagation relies on the assumption that lateral heterogeneities in upper-mantle structure are relatively smooth. In the near future, we shall extend our numerical-adjoint method approach to 3-D earth models, where the only limit to possible Earth's complexity resides in the accuracy of its numerical discretization, and it will be easier to account for the specific geometry of seismic sources—neglected here—and definitions of phase anomaly more realistic than simple cross-correlation—employed here as a first approximation. Recently published applications of analytical finite-frequency methods to surface wave tomography (e.g. Zhou *et al.* 2004, 2005) suggest that the 3-D problem overcomes inherent ray theoretical assumptions made when inverting for 2-D phase-velocity maps as in this study; it will take account of depth-dependent radiation patterns for single scatterers. We therefore, expect the 3-D



**Figure 16.** Phase-velocity maps derived from inversions based on two different heterogeneous starting models. In each case the inversion employed sensitivity kernels calculated via the adjoint method for each respective heterogeneous model. In (a), the starting model is the phase-velocity map of Fig. 9, which was based upon Crust-2.0 and an isotropic upper-mantle model (see Section 4.2). In (b), the starting model is the phase-velocity map of Fig. 14(c), which was derived using numerical-adjoint kernels (see Section 4.4.1).

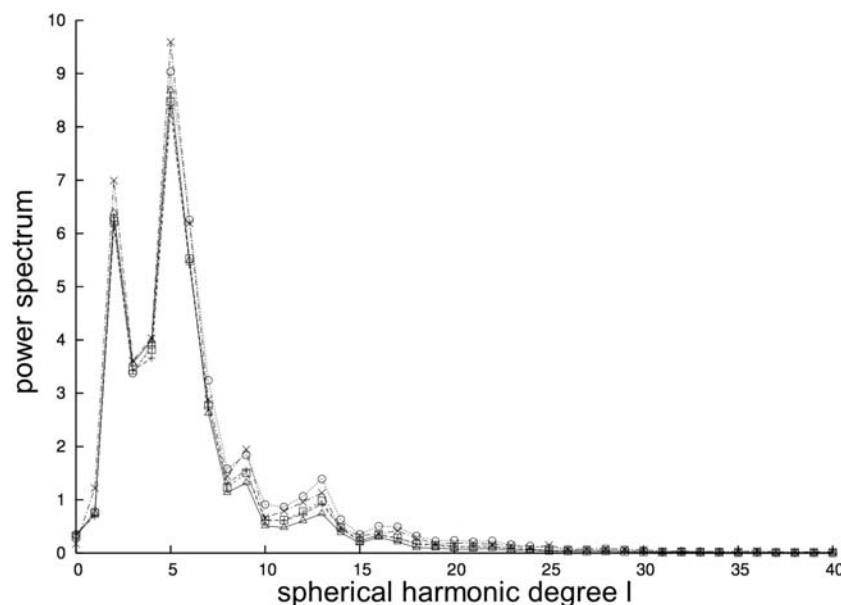
experiment to bring an improvement in data fit and model quality, more significant than what was found here.

## ACKNOWLEDGMENTS

We thank Domenico Giardini for his support and encouragement, Yann Capdeville and Jeroen Tromp for their many insightful comments. We are grateful to Göran Ekström for making his dispersion database available to us. We would also like to thank T. Tanimoto, B. Romanowicz, G. Laske and an anonymous reviewer for critical and constructive comments. Funding for this project is provided by the European Commission's Human Resources and Mobility Program Marie Curie Research Training Network SPICE Contract No. MRTN-CT-2003-504267. Carl Tape's master thesis (Tape 2003) is available at <http://www.gps.caltech.edu/~cartape>

## REFERENCES

- Aki, K. & Richards, P.G., 2002. *Quantitative Seismology*, 2nd edn, University Science Books, Sausalito, CA, USA.
- Baig, A.M. & Dahlen, F.A., 2004. Statistics of traveltimes and amplitudes in random media, *Geophys. J. Int.*, **158**(1), 187–210.
- Baig, A.M., Dahlen, F.A. & Hung, S.H., 2003. Traveltimes of waves in three-dimensional random media, *Geophys. J. Int.*, **153**(2), 467–482.
- Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.*, **81**, F897.
- Baumgardner, J.R. & Frederickson, P.O., 1985. Icosahedral Discretization of the Two-Sphere, *SIAM J. Num. Anal.*, **22**(6), 1107–1115.
- Boschi, L., 2001. Applications of linear inverse theory in modern global seismology, *PhD thesis*, Harvard University.
- Boschi, L., 2006. Global Multi-Resolution Models of Surface Wave Propagation: Comparing Equivalently-Regularized Born- and Ray-Theoretical Solutions, *Geophys. J. Int.*, **167**(1), 238–252.
- Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. geophys. Res.*, **107**(B4), 2059, doi:10.1029/2000JB000059.
- Boschi, L., Ekström, G. & Kustowski, B., 2004. Multiple resolution surface wave tomography: the Mediterranean basin, *Geophys. J. Int.*, **157**, 293–304.
- Boschi, L., Becker, T.W., Soldati, G. & Dziewonski, A.M., 2006. On the relevance of Born theory in global seismic tomography, *Geophys. Res. Lett.*, **33**, L06302, doi:10.1029/2005GL025063.



**Figure 17.** Harmonic spectra of models from Figs 14 and 16. The power spectra are taken from the phase-velocity maps shown in Fig. 14, which were derived from a homogeneous starting model by ray theory (solid line, triangle), analytical Born theory (dashed line, pluses) and numerical-adjoint kernels (short-dashed line, squares), as well as the phase-velocity maps shown in Fig. 16, which were derived from the heterogeneous model shown in Fig. 9 (dash-dotted line, crosses) and Fig. 14(c) (dotted line, circles) employing numerical-adjoint kernels.

Boschi, L., Ampuero, J.-P., Peter, D., Mai, P.M., Soldati, G. & Giardini, D., 2007. Petascale computing and resolution in global seismic tomography, *Phys. Earth planet. Inter.*, in press, doi:10.1016/j.pepi.2007.02.011.

Bunge, H.-P. & Tromp, J., 2003. Supercomputing moves to universities and makes possible new ways to organize computational research, *EOS, Trans. Am. geophys. Un.*, **84**, 30–33.

Capdeville, Y., 2005. An efficient Born normal mode method to compute sensitivity kernels and synthetic seismograms in the Earth, *Geophys. J. Int.*, **163**, 639–646.

Capdeville, Y., Chaljub, E., Vilotte, J.-P. & Montagner, J.-P., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global Earth models, *Geophys. J. Int.*, **152**, 34–67.

Carannante, S. & Boschi, L., 2005. Databases of surface wave dispersion, *Ann. Geophys.*, **48**, 945–955.

Cui, J. & Freeden, W., 1997. Equidistribution on the sphere, *SIAM J. Sci. Comp.*, **18**(2), 595–609.

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

Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.*, **141**(1), 157–174.

de Hoop, M.V. & van der Hilst, R.D., 2005. On sensitivity kernels for ‘wave-equation’ transmission tomography, *Geophys. J. Int.*, **160**, 621–633.

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

Ekström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.*, **102**, 8137–8157.

Favier, N., Chevrot, S. & Komatitsch, D., 2004. Near-field influence on shear wave splitting and traveltime sensitivity kernels, *Geophys. J. Int.*, **156**, 467–482.

Friederich, W., 1999. Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering, *Geophys. J. Int.*, **136**, 180–204.

Heikes, R. & Randall, D.A., 1995a. Numerical integration of the shallow-water equations on a twisted icosahedral grid.1. Basic design and results of tests, *Mon. Wea. Rev.*, **123**, 1862–1880.

Heikes, R. & Randall, D.A., 1995b. Numerical integration of the shallow-water equations on a twisted icosahedral grid.2. A detailed description of the grid and an analysis of numerical accuracy, *Mon. Wea. Rev.*, **123**, 1881–1887.

Hudson, J.A. & Heritage, J.R., 1981. The use of the Born approximation in seismic scattering problems, *Geophys. J. R. astr. Soc.*, **66**, 221–240.

Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.*, **141**, 175–203.

Hung, S.H., Dahlen, F.A. & Nolet, G., 2001. Wavefront healing: a banana-doughnut perspective, *Geophys. J. Int.*, **146**, 289–312.

Kennett, B.L.N., 1972. Seismic waves in laterally inhomogeneous media, *Geophys. J. R. astr. Soc.*, **27**, 301–325.

Komatitsch, D., Ritsema, J. & Tromp, J., 2002. The spectral-element method, Beowulf computing and global seismology, *Science*, **298**, 1737–1742.

Li, X.D. & Tanimoto, T., 1993. Waveforms of long-period body waves in a slightly aspherical Earth model, *Geophys. J. Int.*, **112**(1), 92–102.

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

Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.*, **137**, 805–815.

Moczo, P., Kristek, J. & Halada, L., 2004. *The Finite-Difference Method for Seismologists: An Introduction.*, Comenius University, Bratislava.

Nolet, G. & Dahlen, F.A., 2000. Wave front healing and the evolution of seismic delay times, *J. geophys. Res.*, **105**, 19 043–19 054.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in FORTRAN: the Art of Scientific Computing*, 2nd edn, Vol. xxvi, pp. 963, Cambridge University Press, Cambridge [England]; New York, NY, USA.

Randall, D.A., Ringler, T.D., Heikes, R.P., Jones, P. & Baumgardner, J., 2002. Climate modeling with spherical geodesic grids, *Comput. Sci. Eng.*, **4**, 32–41.

Sadourny, R., Arakawa, A. & Mintz, Y., 1968. Integration of nondivergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere, *Mon. Wea. Rev.*, **96**, 351–356.

Saff, E.B. & Kuijlaars, A.B.J., 1997. Distributing many points on a sphere, *Math. Intell.*, **19**(1), 5–11.

Smith, J.O. III. & Serra, X., 1987. PARSHL: An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation, *Proceedings of the International Computer Music Conference (ICMC-87, Tokyo)*, Computer Music Association.

- Snieder, R., 1987. Surface wave scattering theory, with applications to forward and inverse problems in seismology, *PhD thesis*, Utrecht University, the Netherlands.
- Snieder, R. & Nolet, G., 1987. Linearized scattering of surface waves on a spherical Earth, *J. Geophys.*, **61**, 55–63.
- Spetzler, J. & Snieder, R., 2001. The effect of small-scale heterogeneity on the arrival time of waves, *Geophys. J. Int.*, **145**, 786–796.
- Spetzler, J., Trampert, J. & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.*, **149**, 755–767.
- Strohmaier, E., Dongarra, J.J., Meuer, H.W. & Simon, H.D., 2005. Recent trends in the marketplace of high performance computing, *Parallel Comput.*, **31**, 261–273.
- Stuhne, G.R. & Peltier, W.R., 1996. Vortex erosion and amalgamation in a new model of large scale flow on the sphere, *J. Comput. Phys.*, **128**, 58–81.
- Sword, C.H., Claerbout, J.F. & Sleep, N.H., 1986. Finite-element propagation of acoustic waves on a spherical shell, *Stanford Exploration Project (SEP)*, **50**, 43–78.
- Talagrand, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.
- Tanimoto, T., 1990. Modelling curved surface wave paths: membrane surface wave synthetics, *Geophys. J. Int.*, **102**, 89–100.
- Tanimoto, T., 2003. Geometrical approach to surface wave finite frequency effects, *Geophys. Res. Lett.*, **30**, 1993, doi:10.1029/2003GL017475.
- Tape, C.H., 2003. Waves on a Spherical Membrane, *M.Sc. thesis*, University of Oxford, UK.
- Tape, C., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves, *Geophys. J. Int.*, **168**(3), 1105–1129.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266.
- Tong, J., Dahlen, F.A., Nolet, G. & Marquering, H., 1998. Diffraction effects upon finite-frequency travel times: a simple 2-D example, *Geophys. Res. Lett.*, **25**, 1983–1986.
- Tromp, J. & Dahlen, F.A., 1993. Variational-principles for surface wave propagation on a laterally heterogeneous Earth.3. Potential representation, *Geophys. J. Int.*, **112**, 195–209.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
- Tsuboi, S., Komatitsch, D., Ji, C. & Tromp, J., 2003. Broadband modeling of the 2002 Denali fault earthquake on the Earth Simulator, *Phys. Earth planet. Inter.*, **139**, 305–312.
- Wang, Z. & Dahlen, F.A., 1995. Spherical-spline parameterization of 3-dimensional Earth models, *Geophys. Res. Lett.*, **22**, 3099–3102.
- Wielandt, E., 1987. On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography*, ed. Nolet, G., Reidel, Dordrecht.
- Williamson, D.L., 1968. Integration of barotropic vorticity equation on a spherical geodesic grid, *Tellus*, **20**, 642–653.
- Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized Earth model, *Geophys. J. R. astr. Soc.*, **68**, 653–673.
- Wu, R. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics*, **50**, 582–595.
- Yoshizawa, K. & Kennett, B.L.N., 2005. Sensitivity kernels for finite-frequency surface waves, *Geophys. J. Int.*, **162**, 910–926.
- Zhao, L., Jordan, T.H. & Chapman, C.H., 2000. Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int.*, **141**, 558–576.
- Zhou, Y., Dahlen, F.A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, **158**, 142–168.
- Zhou, Y., Dahlen, F.A., Nolet, G. & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, **163**, 1087–1111.