

Seismic waveform sensitivity to global boundary topography

Andrea Colombi, Tarje Nissen-Meyer, Lapo Boschi and Domenico Giardini

Institut für Geophysik, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland. E-mail: colombia@ethz.ch

Accepted 2012 August 23. Received 2012 August 23; in original form 2012 January 8

SUMMARY

We investigate the implications of lateral variations in the topography of global seismic discontinuities, in the framework of high-resolution forward modelling and seismic imaging. We run 3-D wave-propagation simulations accurate at periods of 10 s and longer, with Earth models including core–mantle boundary topography anomalies of ~ 1000 km spatial wavelength and up to 10 km height. We obtain very different waveform signatures for PcP (reflected) and P_{diff} (diffracted) phases, supporting the theoretical expectation that the latter are sensitive primarily to large-scale structure, whereas the former only to small scale, where large and small are relative to the frequency. PcP at 10 s seems to be well suited to map such a small-scale perturbation, whereas P_{diff} at the same frequency carries faint signatures that do not allow any tomographic reconstruction. Only at higher frequency, the signature becomes stronger. We present a new algorithm to compute sensitivity kernels relating seismic traveltimes (measured by cross-correlation of observed and theoretical seismograms) to the topography of seismic discontinuities at any depth in the Earth using full 3-D wave propagation. Calculation of accurate finite-frequency sensitivity kernels is notoriously expensive, but we reduce computational costs drastically by limiting ourselves to spherically symmetric reference models, and exploiting the axial symmetry of the resulting propagating wavefield that collapses to a 2-D numerical domain. We compute and analyse a suite of kernels for upper and lower mantle discontinuities that can be used for finite-frequency waveform inversion. The PcP and P_{diff} sensitivity footprints are in good agreement with the result obtained cross-correlating perturbed and unperturbed seismogram, validating our approach against full 3-D modelling to invert for such structures.

Key words: Body waves; Computational seismology; Mantle discontinuities; Seismic tomography; Topography; Wave propagation.

1 INTRODUCTION

Since the pioneering works of Oldham (1906), Gutenberg (1914) and Lehmann (1936), among others, mapping the depth and laterally varying topography of seismic discontinuities has been integral to progress in understanding our planet's structure and dynamics. Later research has focused on upper-mantle discontinuities, which are sampled by a relatively large quantity of high-quality seismic data. The 220, 410, 520 and 660 km discontinuities have been studied in detail since the 1960s and are partly features of PREM (Dziewonski & Anderson 1981). The '660' in particular characterizes the bottom end of the transition zone, that is, the boundary separating upper- and lower-mantle subject to important mineral phase transitions (Ita & Stixrude 1992).

More recent studies attempted to map the actual topography of the discontinuities mainly using SS and PP precursor differential traveltimes, see Deuss (2009) for a review. Discontinuity topography is coupled with mantle structure via convective flow (Gu *et al.* 1998; Flanagan & Shearer 1999) so that the tomography maps

can serve as constraints for geodynamics models, including the geometry of subducting slabs and presumed mantle plumes. Seismic studies of discontinuity topography have thus far been based on the ray-theory (infinite frequency) approximation; only Lawrence & Shearer (2008) made use of finite-frequency traveltimes kernels as provided by Dahlen (2005). Waveform inversions have not yet been attempted at the global scale.

The strongest discontinuity in the Earth's interior is the core–mantle boundary (CMB), separating the solid mantle from the fluid outer core. Its topography is very likely related both to the thermal/compositional/viscosity structure (and associated convection) of the mantle (Forte *et al.* 1995; Soldati *et al.* 2012), and to the properties of the outer core, where vigorous convection is believed to generate the Earth's magnetic dynamo (Jackson *et al.* 1993). A number of authors, starting with Morelli & Dziewonski (1987), mapped CMB topography based on compressional-wave traveltimes, whereas others (e.g. Ishii & Tromp 2001; Koelemeijer *et al.* 2012) inverted observations of eigenfrequency splitting, focusing on normal modes sensitive to the CMB. Most ray-theory authors

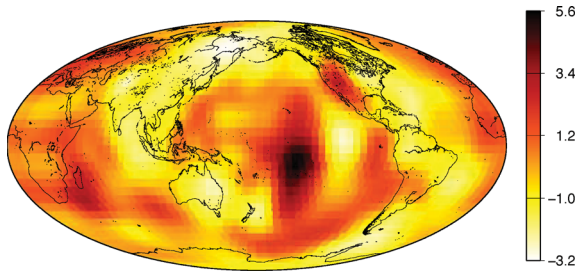


Figure 1. Map of CMB topography obtained in Soldati *et al.* (2012). Note the smooth lateral variation. The topography is measured in kilometre.

find a similar degree-2 pattern with peak-to-peak topography of a few kilometres (see Fig. 1 for an example), but shorter-wavelength structure is more difficult to constrain. Several studies (e.g. Boschi & Dziewonski 1999; Vasco *et al.* 1999; Soldati *et al.* 2003) point to a discrepancy in CMB structure as mapped by core-refracted (various branches of the *PKP* phase) versus core-reflected seismic waves, which casts some doubt on the validity of CMB maps derived from those data. Soldati *et al.* (2012) show that the discrepancy can be reduced if one requires the inverse-problem solution for CMB topography to be coupled with seismic structure in the mantle according to the theory of Forte *et al.* (1995).

With this study, we present a wave-based formulation of the seismic inverse problem associated with discontinuity topography, whose future applications to global seismic databases should help to enhance the resolution of our CMB maps. As noted, for example, by Dahlen (2004) and Montelli *et al.* (2004), ray and finite-frequency inversions may give similar results when applied to *P* or *PcP* data, but stronger differences are to be expected if the more complex, core-refracted (*PKP*) or diffracted (P_{diff} , S_{diff}) waves are taken into account. Our formulation is based on first-order Born perturbation theory (Nissen-Meyer *et al.* 2007a) coupled with the adjoint method (Geller & Hara 1993; Tarantola 2005; Tromp *et al.* 2005; Peter *et al.* 2007). Boundary sensitivity kernels presented here can naturally be applied in a joint volumetric discontinuity inversion.

Although algorithms for joint waveform inversion of both volumetric and discontinuity heterogeneity are still under development, independent finite-frequency inversions for volumetric or discontinuity structures have been attempted. Much attention was devoted to volumetric inversions (Montelli *et al.* 2004; Fichtner *et al.* 2009; Tape *et al.* 2009), while finite-frequency discontinuity topography has been the subject of very few (Dahlen 2005; Takeuchi 2005; Lawrence & Shearer 2008) studies. Discontinuity imaging is, in fact, more difficult than volumetric tomography: although the latter aims at constraining structures that presumably extend for hundreds of kilometre in all directions, the topography of known mantle discontinuities never exceeds a few kilometres radially and thus requires very specific source-receiver layout.

A methodology for joint volumetric-discontinuity finite-frequency inversions of seismic data is proposed by Mora (1989) in the context of exploration seismology. The treatment that we shall develop here is similar to that of Mora (1989); both allowing to enhance imaging resolution via the combination of different, reflected, diffracted and transmitted seismic phases. Our method, however, is designed to be applicable under very different underlying hypotheses. Namely, compared to the typical wavelengths of exploration seismology, at the global scale, the average depths of global discontinuities are known rather well and topography is expected to be a weak perturbation to these radial discontinuities; volumetric

heterogeneity is not as strong in the deep Earth; the source-station distribution is much less uniform, and cannot be controlled.

In this work, we apply the axi-symmetric methodology first developed by Nissen-Meyer *et al.* (2007a, 2008) to study the effects of boundary topography on waveforms and to determine boundary sensitivity kernels for a suite of different seismic phases. Important issues that we study in detail are (i) the physics of global seismic wave propagation in the presence of undulating interfaces (Section 2), in particular at the CMB (Section 2.5), and (ii) our strategy to reduce the otherwise immense computational costs (Sections 3.1 and 3.4) for these inverse problems. This is followed by a gallery of traveltimes boundary sensitivity kernels (Section 4), whereas an appendix collects further details on the methodology.

2 THE FORWARD PROBLEM: EFFECTS OF BOUNDARY PERTURBATION ON SEISMOGRAMS

2.1 Global wave propagation: notation and analytic background

In our convention, vectors are denoted as boldface lower-case letters (\mathbf{u} , \mathbf{v} , ...) and tensors as uppercase letters (\mathbf{T} , \mathbf{E} , ...). Vectors and tensors are real-valued functions $\in \mathbb{R}^3$. For further reference about vector conventions and operations used in this paper, refer to Dahlen & Tromp (1998).

The physical quantities have the usual meaning of linear elasticity problems formulated from a Lagrangian perspective. We define three-component vectors $\mathbf{u}(\mathbf{x}, t)$ as the displacement at point \mathbf{x} at time t , measured in metres. The first and second derivatives with respect to t are, respectively, velocity: $\dot{\mathbf{u}}(\mathbf{x}, t)$, $\partial_t \mathbf{u}(\mathbf{x}, t)$ measured in m/s and acceleration $\ddot{\mathbf{u}}(\mathbf{x}, t)$ or $\partial_t^2 \mathbf{u}(\mathbf{x}, t)$ measured in m/s^2 . Whenever possible, we render the dependence of various physical quantities on \mathbf{x} and t implicit.

The dimensionless strain tensor is defined as the symmetric part of the displacement gradient $\mathbf{E} = \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^T]$. Hooke's relationship then gives the stress tensor: $\mathbf{T} = \mathbf{C} : \mathbf{E}$ measured in Pascal upon the definition of \mathbf{C} , the fourth-order tensor containing at most 21 independent elastic constants. The practical case studies are limited here to isotropy, but our formulation remains valid for general anisotropic media.

At the global scale, the phenomenon of seismic wave propagation is explained well by linear elasticity theory. The amplitude of seismograms is affected significantly by attenuation; their phase only marginally (Komatitsch & Tromp 2002b). The effects of Earth's rotation and gravitation and of the oceans are all negligible at the frequency band (40–100 mHz) and geographic scale length (20° – 180°) that will be considered here (Dahlen & Tromp 1998; Komatitsch & Tromp 2002b), whereas those of the Earth's ellipticity can be accounted for by simply applying a linear correction on the data as shown by Dziewonski & Gilbert (1976).

Under these assumptions and after defining an Earth model \mathbf{m} , wave propagation in the whole volume of the Earth Ω is described by the system:

$$\partial_t^2 \mathbf{u}(\mathbf{x}, t) + \nabla \cdot \mathbf{T}(\mathbf{x}, t, \mathbf{m}) = \mathbf{f}(\mathbf{x}, t) \quad \text{in } \Omega, \quad (1a)$$

$$\mathbf{T}(\mathbf{x}, t, \mathbf{m}) = \mathbf{C}(\mathbf{m}) : \mathbf{E}(\mathbf{x}, t) \quad \text{in } \Omega, \quad (1b)$$

$$\hat{\mathbf{n}} \cdot \mathbf{T} = 0 \quad \text{on } \partial \Omega, \quad (1c)$$

$$\text{Interface Condition} \quad \text{on } \Sigma \quad (1d)$$

$$\mathbf{u}(\mathbf{x}, 0) = 0 \quad \partial_t \mathbf{u}(\mathbf{x}, t) = 0, \quad (1e)$$

where $\hat{\mathbf{n}}$ is the normal vector and $\partial\Omega$ the surface of the Earth. Eq. (1a) expresses conservation of momentum, with the forcing term $\mathbf{f}(\mathbf{x}, t) = \mathbf{M}\dot{m}(t)$ a function of the focal mechanism (through the moment tensor \mathbf{M}) and of the source-time function $\dot{m}(t)$ of the earthquake. Eq. (1b) is known as the constitutive relationship that is defined by the selected Earth's reference model and in seismology is often taken to coincide with Hooke's law. The boundary conditions (1c) require the outer surface of the Earth to be free of tractions. Interface condition (1d) links regions with different physics or discontinuous properties. The solution of a differential equation like (1) requires the initial conditions (1e). In our case, the Earth is considered to be at rest at $t = 0$, prior to occurrence of an indigenous seismic event represented by the forcing term $\mathbf{f}(\mathbf{x}, t)$.

When eq. (1) is used to model the propagation of realistic seismic waves in the Earth, it is crucial to choose the source-time function $\dot{m}(t)$ appropriately. We assume $\dot{m}(t)$ to be Gaussian; this is common practice in seismology, signal processing and image processing, and be written as: $m(t) = \exp(-\gamma(t + T_0 - \tau)^2)$, where T_0 is the dominant period, τ the shift from $t = 0$ and γ the shape factor.

Improving the level of detail and elaborating models that are consistent with geological, geodynamic and geomagnetic constraints is the main goal of seismic tomography. Although most authors currently agree on the long-wavelength ($\gg 1000$ km) pattern of mantle structure (e.g. Becker & Boschi 2002), no agreement has been achieved on the nature of relatively small-scale tomographic features interpreted as, for example, mantle plumes or subducting slabs. Spherically symmetric models such as PREM (Dziewonski & Anderson 1981) explain up to 90 per cent of traveltime observations. This is justified by the fact that very strong lateral variations are expected only in a small portion of the total volume of the Earth: the highly heterogeneous crust, and the area surrounding the CMB, while the remaining volume is characterized by heterogeneity of small amplitude (~ 1 per cent for P and ~ 2 per cent S velocity).

2.2 Boundary conditions

Regardless of the reference model considered, all global tomographic models share some similarities. Spherically symmetric reference models are defined as the union of spherical-shells domain $\Omega = \Omega_1 \cup \Omega_2 \cup \dots$, interfaces Σ linking different domains, and free surface $\partial\Omega$ that bounds the outer domain where external forces are applied. Domains are all solid except for the outer core. Oceans are directly modelled at local scale, while for global studies, they can be treated as external forces (Komatitsch & Tromp 2002b). Continuity between domains is guaranteed by the transmission condition defined at various interfaces. Depending on the nature of the two domains, only certain quantities may be transmitted.

The jump of a quantity $[\mathbf{u}]_+^+$ across an interface is obtained taking the difference of the value starting from the outer (+) to the inner (−) boundary. The transmission conditions at interface (1d) then read

$$[\mathbf{u}]_+^+ = 0 \quad \text{and} \quad [\mathbf{T}]_+^+ = 0 \quad \text{on} \quad \Sigma_{SS}, \quad (2a)$$

$$[\mathbf{u} \cdot \hat{\mathbf{n}}]_+^+ = 0 \quad \text{and} \quad [\mathbf{T} \cdot \hat{\mathbf{n}}]_+^+ = 0 \quad \text{on} \quad \Sigma_{SF}, \quad (2b)$$

where Σ_{SS} and Σ_{SF} denote the union of all solid–solid and all solid–fluid interfaces, respectively. Eq. (2b) is also valid for time derivatives of enclosed quantities. Material properties vary smoothly within each layer. In the case of PREM, they are analyti-

cally defined as 1-D polynomial functions (Dziewonski & Anderson 1981).

The solid–solid condition is naturally embedded inside the elastodynamic system. When equations are solved via finite-element methods, these conditions are fundamental part of the discretization procedure (hence, the jump of the quantities across the interface is entirely due to the discontinuity in the material property).

Solid–fluid boundaries, on the other hand, involve a change in the physics of wave propagation, as no shear waves can exist in a fluid medium. The elastodynamic system becomes a scalar problem formulated in terms of a potential function. The two systems are coupled together, for instance as in Nissen-Meyer *et al.* (2008).

2.3 Computational cost of global-scale wave-propagation modelling

Although high-resolution 3-D models of the entire Earth's mantle (e.g. Bijwaard *et al.* 1998; Montelli *et al.* 2006; Li *et al.* 2008; Obayashi *et al.* 2009; Ritsema *et al.* 2010; Simmons *et al.* 2010; Della Mora *et al.* 2011) are available, their direct application is bounded by computational resources, particularly if notoriously expensive waveform inversions were to be conducted. Fig. 2 shows that computation time for a global 3-D simulation (SPECFEM3D GLOBE of Komatitsch & Tromp 2002a) is roughly inversely proportional to the cube of the shortest resolved period T_0 . The axisymmetric spectral element code by Nissen-Meyer *et al.* (2008), hereafter named AXISEM, models exact 3-D wave propagation in a spherically symmetric Earth by solving the equations of motion numerically on a 2-D disc; the size of the mesh, and, consequently, the computational costs are massively reduced, with runtime inversely proportional to the square of T_0 .

We infer that with the current typically available computational infrastructure, it is convenient, if not necessary, to ground global finite-frequency tomography (and the numerical computation of sensitivity kernels) to spherically symmetric Earth models, particularly if full-waveform inversions are to be conducted. Sensitivity kernels throughout this study are calculated by means of AXISEM unless noted otherwise.

2.4 Ground-truth synthetic database via 3-D modelling

The adjoint technique is at the core of finite-frequency full waveform tomography for computational reasons, relies on first-order

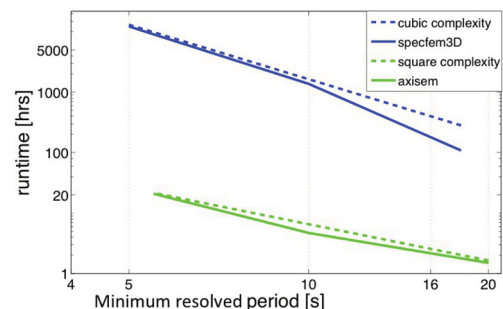


Figure 2. Computation time of SPECFEM3D GLOBE (blue lines) and AXISEM (green) as a function of the shortest resolved period, to model a 25-min long seismogram based on PREM. The time is meant cumulative across all CPUs. Experimental results (solid lines) are compared with the theoretically expected scaling of $\mathcal{O}((\frac{1}{T_0})^3)$ for SPECFEM3D GLOBE (3-D integration of the equation of motion) and $\mathcal{O}((\frac{1}{T_0})^2)$ for AXISEM (2-D integration). The number of CPU spans from 2 to 8 for AXISEM, while from 256 to 1944 for SPECFEM3D.

perturbation theory, that is, a linearization that might or might not be valid depending on how far the starting model is from the true one. Its accuracy and applicability depend on data type and frequency, their error statistics, the chosen misfit function, the sparsity of data coverage, the wavelength of the anomaly we wish to resolve, the signal-to-noise ratio and inversion parameterization. Before addressing the inverse problem, forward simulations of core-sensitive seismic phases are useful to evaluate the effects of the mentioned parameters. We construct a synthetic database based on numerical integration of the equation of motion via SPEC-FEM3D GLOBE (Komatitsch & Tromp 2002a). Unlike AXISEM, SPEC-FEM3D GLOBE accommodates 3-D velocity and/or topography heterogeneity; comparing the results of SPEC-FEM3D GLOBE with those of first-order perturbation theory allows us to quantify the accuracy of the latter for different wavelength-to-topography length scales and seismic phases and determine settings which optimally highlight topography.

Our experimental setup is illustrated in Fig. 3. To isolate the effects of CMB topography, we assume a spherically symmetric Earth model (PREM). For the sake of simplicity, we neglect the effects of surface topography, crustal heterogeneity, ellipticity and gravitation (relevant at $T_0 \geq 100$ s). We simulate explosive events located around a single region of anomalous, smooth CMB topography. Seismograms calculated with SPEC-FEM3D GLOBE are valid down to 10 s period. After the simulation, we can filter the modelled traces to isolate any desired frequency band, with no need to repeat the same simulation for different frequencies.

We utilize different source-station distributions depending on the phase that we wish to analyse. Fig. 4a shows a dense, relatively low-aperture receiver network designed for the small-aperture phase PcP : a $60^\circ \times 60^\circ$ area around the centre of the anomalous CMB area is covered by receivers every 1° . The P_{diff} phase is naturally characterized by a broader range of possible epicentral distances, and is accordingly analysed on the wider-aperture array ($120^\circ \times 120^\circ$ with 2° spacing) of Fig. 4b. We also experiment with different patterns of CMB topography: closely spaced depression and uplift (Fig. 5b), defined by the colatitude derivative of a Gaussian function, or a broad (Gaussian) uplifted region (Fig. 5a). A section of seismograms due to an event at the bottom-left for the rightmost column of receivers depicted in the event-receiver layout in Fig. 4a is portrayed in Fig. 6.

As anticipated by Fig. 2, these tests are computationally very expensive. To simulate a 25-min-long seismic record accurate down to 10 s period, we need 200 Gbyte of system memory distributed over 486 processors for a wall clock time of about 3 hr. The construction of a (complete) database spanning several combinations of simulation parameters required $\sim 5 \times 10^5$ CPU hours.

2.5 Probing the database for body-wave-CMB interaction

We measure the residual time $\delta t = t_{prem} - t_{topo}$ of a given phase as the cross-correlated time difference between perturbed t_{topo} and unperturbed seismogram t_{prem} . The cross-correlation window is centred at the expected arrival time, calculated with TauP (<http://www.seis.sc.edu/software/TauP/>), and its width is proportional to the frequency content of the seismogram. Seismograms are filtered with a Gaussian function with half-width/period corresponding to the target frequency. We investigate four different frequency bands with periods: 30, 20, 15 and 10 s, filtering seismograms before cross-correlation. We iteratively tune the cross-correlation window's width to optimize its bounds with respect to the anomaly.

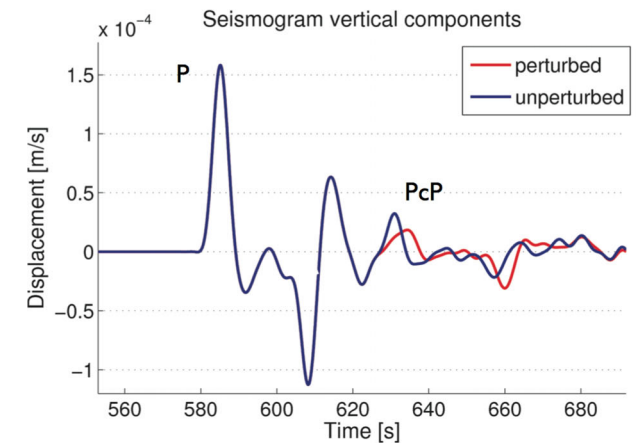
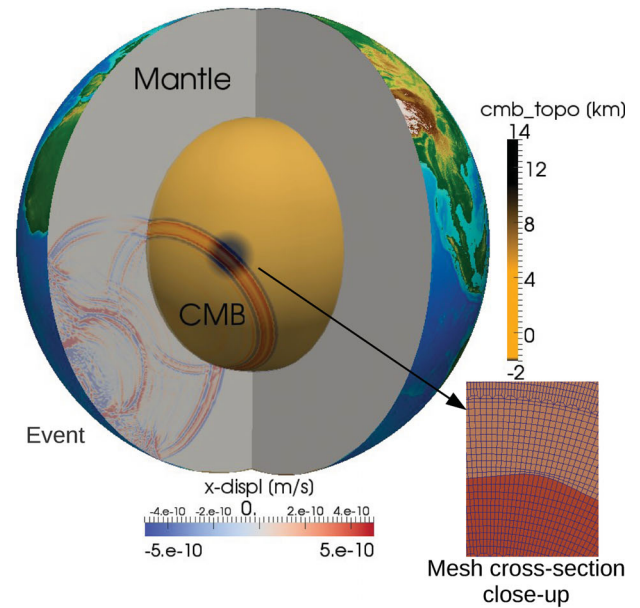


Figure 3. (Top) A section of the Earth showing one of our synthetic CMB topography models: in this case, topography is Gaussian with a maximum height of 11 km and a lateral extension of ~ 800 km. The anomalous CMB area is visible as a dark circle at the centre of the plot. A cross-section of the global mesh taken at the CMB shows the stretching (10 times exaggerated) of mesh elements to accommodate topography. A 100-km-deep explosive event was located at 30° S and 10° E. (Bottom) reference (CMB with constant radius) (blue line) and perturbed (red line) seismograms predicted at a station aligned with source and CMB anomaly, 23° away from the source. The largest misfit corresponds to the PcP arrival time. The perturbed seismogram is obtained after amplifying topography to 70 km, to emphasize its effects.

We note first of all, and show in Fig. 7, that, as predicted by finite-frequency theory, the effect of a CMB topography anomaly on a seismic measurement can be significant even if the associated ray path does not sample the anomaly directly.

2.5.1 PcP synthetics

We show in Fig. 8 a set of cross-correlation-based PcP travel-time anomalies for different source-anomaly-receiver geometries and different CMB topography patterns. The plots define the pattern of residuals as a function of receiver location, with a CMB uplift resulting in earlier PcP arrivals, and a CMB depression in

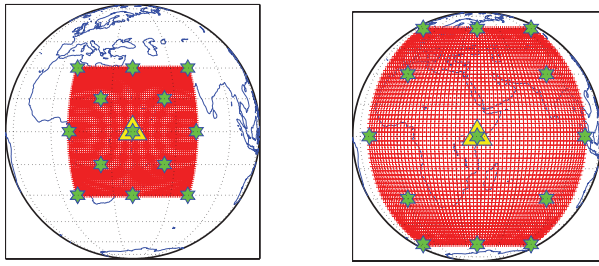
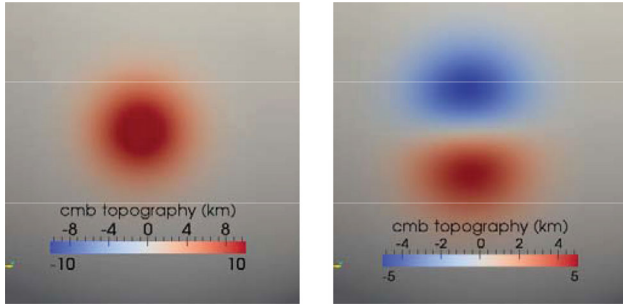


Figure 4. (a) The layout of receivers (red crosses), events (green stars) and topography anomaly (yellow triangle) for low-aperture phases like *PcP*, *ScS* and *PcS*. (b) Same as (a), but for long-epicentral-distance phases like *P_{diff}*, *PKP*.



(a) *Bell-shaped uplift*

(b) *Depression/uplift*

Figure 5. Patterns of synthetic topography models used in this study; the lateral extent of both maps is ~ 800 km.

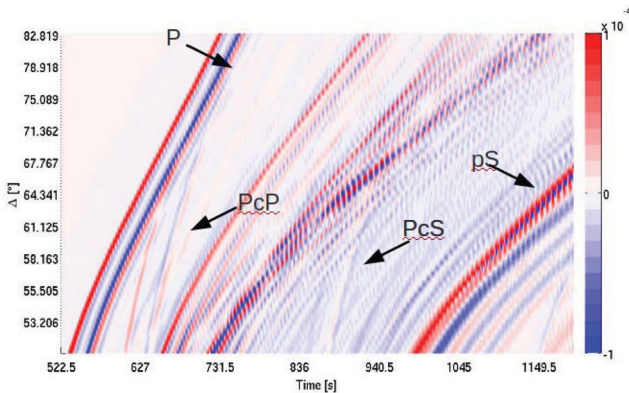


Figure 6. Vertical component section of 61 stations for one of the source-receiver configurations of Fig. 4. Main phase arrival time is also marked. The y -axis shows epicentral distance, and x -axis time in seconds. Colours depend on the magnitude of the displacement in metre.

delayed arrivals. The misfit also depends upon the relative position of the source with respect to the topography anomaly, and phases that bottom (in a ray-theoretical sense) between 20° and 30° away from the source are more strongly affected by CMB topography.

To summarize the cumulative effect of CMB heterogeneity on the entire synthetic database, we introduce the misfit functions

$$L_1 = \sum_{i=1}^N |\delta t_i| \quad \text{and} \quad L_\infty = \max_N |\delta t_i|, \quad (3)$$

with N the total number of source-receiver pairs. The L_∞ norm reflects the maximum residual time present in the database, while the L_1 norm gives an idea of what portion of seismograms (i.e. source-receiver couples) are affected by the anomaly. L_∞ and L_1

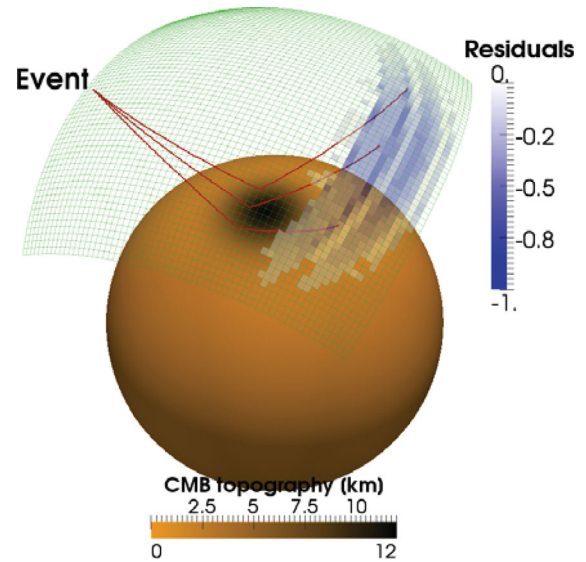


Figure 7. Ray paths for three different stations are plotted along with the measured misfit. The effects of topography on SPEC-FEM3D GLOBE synthetics are not only evident for rays bottoming at the centre of the anomaly, but also for those grazing its sides, for which ray theory predicts no travel-time anomaly.

misfits for two different realizations of the *PcP* synthetic database are shown in Fig. 9. We also compute the ray-theoretical residual

$$\delta t_{\text{ray}} = -2 \frac{\delta d}{r} \sqrt{\left(\frac{r}{v(r)}\right)^2 - p^2} \quad (4)$$

(e.g. Morelli & Dziewonski 1987) (where r is the unperturbed CMB radius, δd its perturbation at the bounce point, $v(r)$ the P -wave speed and p the ray parameter), and show the corresponding L_∞ norm in Fig. 9. For the L_1 norm, ray-theoretical values are not shown as they are out of scale.

The L_∞ norm in Fig. 9 indicates a maximum residual of ~ 1 s, lower than the 1.5 s residual predicted by eq. (4). Errors on these values associated with our automated cross-correlation procedure can vary roughly between 0.1 and 0.2 s. The discrepancy between ray-theory predictions and numerical results grows with increasing period, as is to be expected given that ray theory is strictly valid only at the infinite-frequency limit. Synthetics do not converge to the ray-theoretical value because they are calculated for a perfect vertical reflection that is not visible from synthetics. Hence, the actual upper bound is positioned slightly below.

It is also clear from Fig. 9 that the potential to resolve topography (as well as volumetric) heterogeneity depends strongly on the frequency of the analysed signal: filtering the synthetics around 30 s would essentially obliterate the CMB-topography signature and we should utilize data at ~ 10 s or lower to be able to detect these signatures, that is, keeping the lateral length-scale to wavelength ratio ≥ 5 . Decreasing the lateral extension of anomalous topography does not affect strongly the L_∞ misfit as the height does, indicating that seismic waves are more sensitive to the amplitude of topography rather than its shape. The L_1 misfit shows a lower value as expected with a reduced extension of the topography.

A similar experiment can be conducted upon the CMB topography of Fig. 5b, with positive and negative peaks at ± 6 km. The resulting residual traveltime maps are shown in Fig. 8. The abrupt change in the sign of the residual again reflects the high sensitivity to the anomaly height.

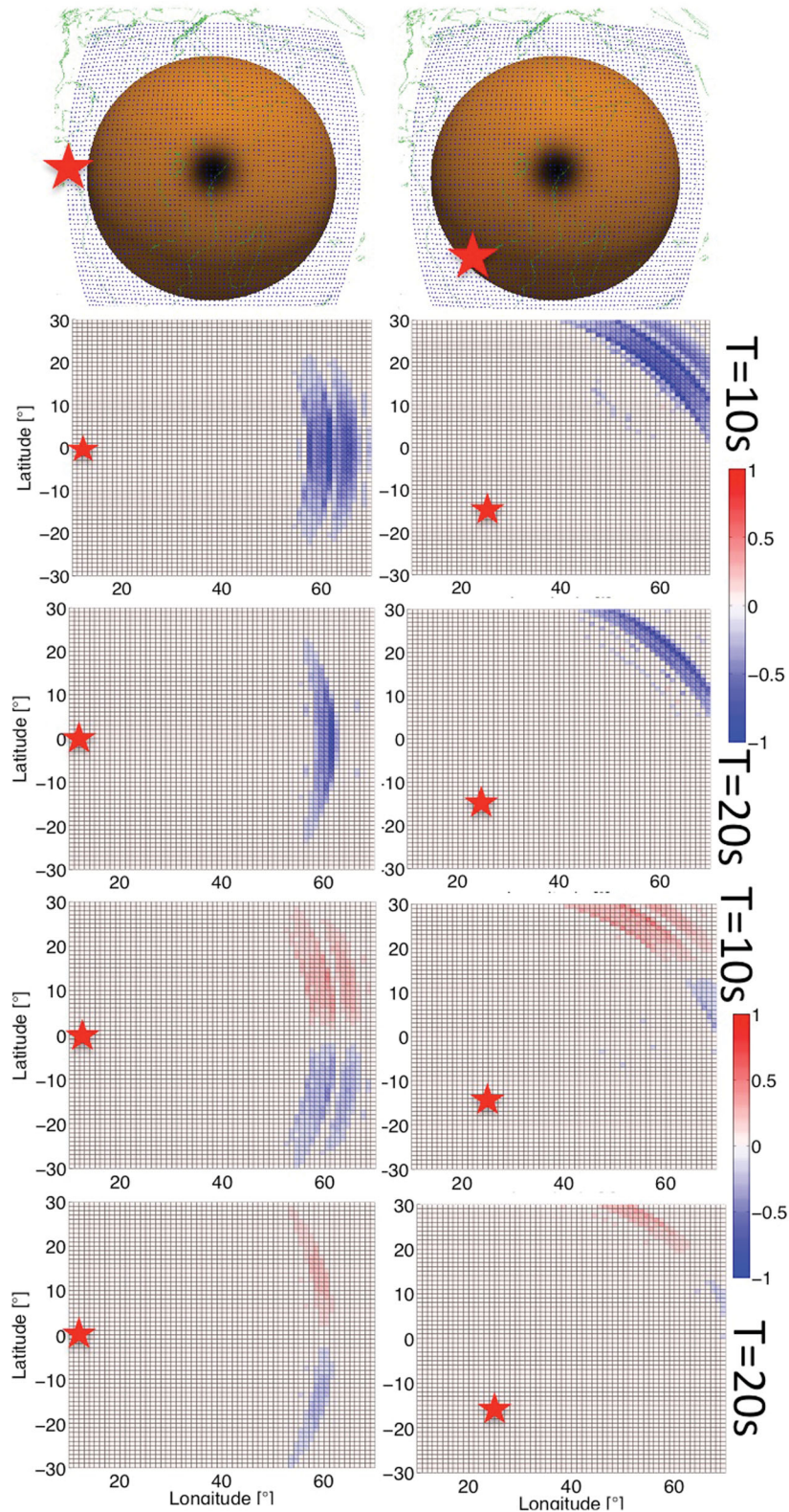


Figure 8. PREM versus PREM + topo residual *PcP* traveltimes plotted at each receiver's location. Two different (left versus right panels) source locations, marked by a red star in the top panels, are simulated. The top four residual maps refer to the topography model of Fig. 5a. Seismograms are filtered around the dominant period (10 or 20 s) specified to the right of each row. The bottom four panels refer to the topography model of Fig. 5b with analogous source position.

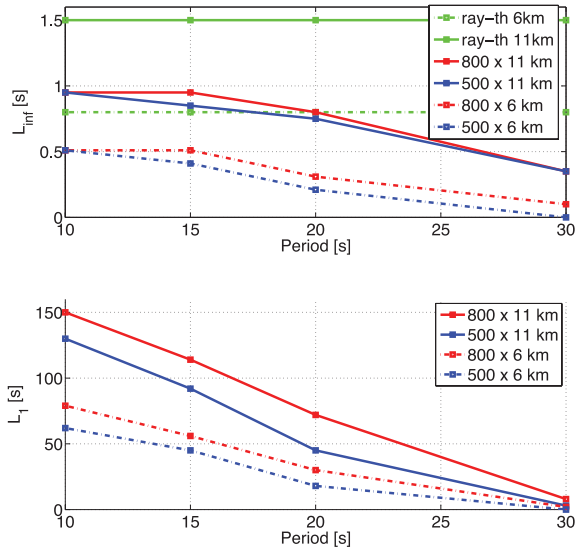


Figure 9. The L_∞ (top) and L_1 (bottom) cumulative misfits between PREM synthetics and two sets of anomalous CMB PcP synthetics, as a function of the dominant period of filtered signal, generated after imposing a Gaussian-shaped topography anomaly with a maximum height of 11 km (continuous) and 6 km (dashed), and a lateral extent of (red) 800 km and (blue) 500 km. The green lines in the top panel show the frequency-independent prediction of ray theory, for the ray path bouncing at the highest topographic peak (L_∞ misfit).

Because of their reflected origin, PcP signals are very sensitive to the height of the topography rather than its lateral extent. If, on the one hand, we can get useful data from reflected phases, on the other hand, we also need to be particularly careful about their spatial coverage, since they might be very strongly affected by topography heterogeneities of very small scale.

2.5.2 P_{diff} synthetics

Fig. 10 shows that even at the relatively short period of 10 s, the assumed CMB anomaly (Fig. 5a) has a very faint effect on travel-times of P_{diff} . Residuals remain small even if the lateral extent of the topography anomaly is doubled. At longer periods (not shown here for brevity), the signal is lost completely. These observations are confirmed by direct inspection of individual seismograms.

3 THE INVERSE PROBLEM FOR TOPOGRAPHY

Full waveform inversion can make complete usage of the residuals computed in the previous section, but expensive computation of the Fréchet derivatives of data with respect to model parameters has to be carried out (fortunately the adjoint methodology comes into help, making this computation feasible).

3.1 Seismic sensitivity kernel foundations

A seismic anomaly δu is linked to model perturbations δv and δd , respectively, volumetric and boundary, through sensitivity kernels \tilde{K}_v and \tilde{K}_d :

$$\delta u(t) = \int_{\Omega} \delta v \tilde{K}_v(\mathbf{x}, t) d\mathbf{x}^3 + \int_{\Sigma} \delta d \tilde{K}_d(\mathbf{x}, t) d\mathbf{x}^2. \quad (5)$$

Here, we just focus on the boundary contribution.

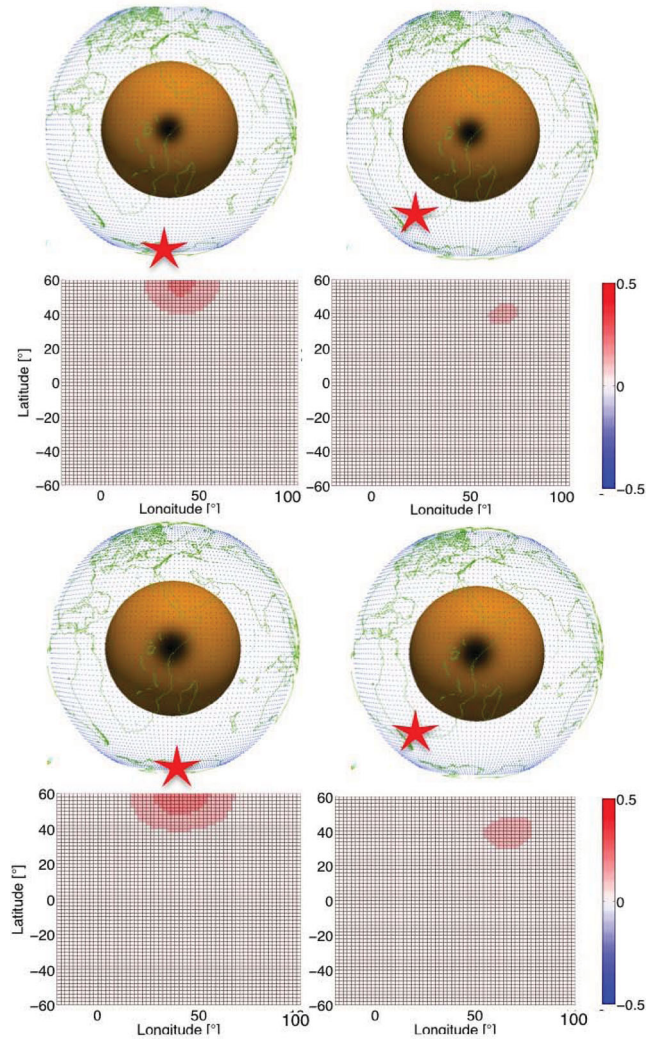


Figure 10. Residual P_{diff} travel-times plotted at each receiver's location, similar to Fig. 8. Two different (left versus right panels) source locations, marked by a red star in the top panels, are simulated. Residuals generated by two Gaussian CMB topography models: (top) 800 km in lateral extent and 11 km in maximum height, and (bottom) 1600 km lateral extent, same height.

Boundary sensitivity kernels in the first-order Born approximation were first obtained analytically by Dahlen (2005), following the scheme first developed for volumetric kernels by Dahlen *et al.* (2000) and Hung *et al.* (2000). The numerical implementation of terms \tilde{K}_v and \tilde{K}_d in eq. (5) is computationally prohibitive for full 3-D wave propagation at high resolution. Computational cost is reduced particularly if, as mentioned above, a 1-D reference model is used leading to the integration of the equations of motion on a 2-D disc, as in Nissen-Meyer *et al.* (2007b, 2008). We follow these studies and complement the corresponding software with an additional module, which provides seismic sensitivity to boundary topography for any of the interfaces included in the selected reference model such as upper-mantle discontinuities and both core boundaries.

3.2 Waveform sensitivity to boundary perturbations

The calculation of boundary sensitivity requires the decomposition of a vector in normal and tangential component: $\mathbf{u} = u \hat{\mathbf{n}} + \mathbf{u}^\Sigma$.

The surface gradient and divergence (Dahlen & Tromp 1998) are accordingly defined:

$$\text{surface gradient of } \mathbf{u} : \quad \nabla^\Sigma \mathbf{u} = \nabla \mathbf{u} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \nabla \mathbf{u}), \quad (6)$$

$$\text{surface divergence of } \mathbf{u} : \quad \nabla^\Sigma \cdot \mathbf{u} = \text{tr}(\nabla^\Sigma \mathbf{u}), \quad (7)$$

where $\text{tr}()$ is the tensor-trace operator. For notational brevity, we denote with $(*)$ the convolution between two time-dependent quantities (written in the time domain)

$$\mathbf{g}(t) * f(t) = \int_0^t \mathbf{g}(\tau) f(t - \tau) d\tau. \quad (8)$$

The convolution between vector (tensor) quantities is a scalar, and involves (double) contraction, that is,

$$\mathbf{g}(t) * \mathbf{h}(t) = \int_0^t \mathbf{g}(\tau) \cdot \mathbf{h}(t - \tau) d\tau \quad \text{vectors}, \quad (9)$$

$$\mathbf{G}(t) * \mathbf{H}(t) = \int_0^t \mathbf{G}(\tau) : \mathbf{H}(t - \tau) d\tau \quad \text{tensors}, \quad (10)$$

The model parameter vector \mathbf{m} is defined as the input of the forward problem, and the unknown of the inverse problem. \mathbf{m} can, in principle, include all the elastic parameters of the medium of propagation including shear and compressional velocity, their anisotropy, density, as well as source mechanism and location. In the framework of first-order perturbation theory, it is more convenient to consider its variation $\delta \mathbf{m}$. For brevity, we assume that we solve an inverse problem for boundary perturbation alone: $\delta \mathbf{m} = [\delta d]$. The boundary perturbation δd must be small enough for the first-order perturbation theory to be a valid approximation. Our formulation is valid for any discontinuity defined by the reference model.

The boundary sensitivity kernel has a more complicated form than the volumetric type (Nissen-Meyer *et al.* 2007a). The expression calculated for the frequency domain by Dahlen (2005) may be computed for the time domain and, making use of (10) to avoid clutter, we can write:

$$\tilde{K}_d(\mathbf{x}, t) = \underbrace{\left[\rho \partial_t \vec{\mathbf{u}} * \partial_r \vec{\mathbf{u}} + \vec{\mathbf{T}} * \vec{\mathbf{E}} - \vec{\mathbf{t}} * \nabla_{\hat{\mathbf{n}}} \vec{\mathbf{u}} - \vec{\mathbf{t}} * \nabla_{\hat{\mathbf{n}}} \vec{\mathbf{u}} \right]}_{\text{general interface}} \quad (11)$$

$$\underbrace{\left[-\hat{\mathbf{n}} \cdot \vec{\mathbf{t}} * \nabla^\Sigma \cdot \vec{\mathbf{u}} - \hat{\mathbf{n}} \cdot \vec{\mathbf{t}} * \nabla^\Sigma \cdot \vec{\mathbf{u}} - \vec{\mathbf{u}} * \nabla^\Sigma (\hat{\mathbf{n}} \cdot \vec{\mathbf{t}}) - \vec{\mathbf{u}} * \nabla^\Sigma (\hat{\mathbf{n}} \cdot \vec{\mathbf{t}}) \right]^+}_{\text{fluid/solid interface}}, \quad (12)$$

where $\tilde{K}_d(\mathbf{x}, t)$ is the time-dependent sensitivity of the waveform with respect to the boundary topography and $\vec{\mathbf{t}}$ the traction acting on the interface. The solid–fluid term arises from interface conditions (2b). We denote with a right-pointing arrow as in $\vec{\mathbf{u}}$, the ‘forward field’, that is, the regular propagating field from instant $t = 0$ to $t = T$ emanating from the earthquake location \mathbf{x}_s , while the backward field, denoted by a left-pointing arrow, is a time-reversed field going from negative time $t = -T$ to $t = 0$, emanating from the receiver position \mathbf{x}_r . Derived quantities such as $\vec{\mathbf{E}} = \frac{1}{2}[\nabla \vec{\mathbf{u}} + (\nabla \vec{\mathbf{u}})^T]$ and $\vec{\mathbf{T}} = \mathbf{C} : \vec{\mathbf{E}}$ are calculated in the same way in both forward and backward fields.

In the adjoint framework, the definition of misfit modifies the expression for $\vec{\mathbf{f}}$ and hence the form of the backward field (Tromp *et al.* 2005; Peter *et al.* 2007; Fichtner *et al.* 2009).

The forward field $\vec{\mathbf{u}}$ is now defined as the solution of eq. (1) for a moment tensor source centred at \mathbf{x}_s ,

$$\vec{\mathbf{f}}(\mathbf{x}, t) = \mathbf{M}\dot{m}(t)\delta(\mathbf{x} - \mathbf{x}_s), \quad (13)$$

where $\mathbf{M}\dot{m}(t)$ has been defined in Section 2.1. The backward field is instead the time reversed solution to (1) for an impulsive forcing term

$$\vec{\mathbf{f}}(\mathbf{x}, t) = \hat{\mathbf{x}}_r \delta(\mathbf{x} - \mathbf{x}_r) \delta(T - t), \quad (14)$$

that is, the Green’s function associated with (1) and the receiver locations, \mathbf{x}_r . The vector $\hat{\mathbf{x}}_r$ points in the same direction of the seismogram component used to measure δu . For instance, if only vertical components are used to compute \tilde{K}_d , then $\vec{\mathbf{f}}$ is a vertical single force and the backward field the numerical solution to the heterogeneous extension of Lamb’s problem (Lamb 1904). Note that only one simulation is necessary for all distances and configurations.

3.3 Traveltime sensitivity to boundary perturbations

The most general expression relating a seismic traveltime anomaly observation to discontinuity topography in the first-order approximation is implicitly found from eq. (5) after dropping the volumetric term,

$$\delta t = \int_\Sigma \delta d(x) K_d(\mathbf{x}) d\mathbf{x}^2, \quad (15)$$

where the sensitivity kernel $K_d(\mathbf{x})$ is naturally not the same as that of eq. (5). Note that (11) is independent from the measured misfit. Only after having defined an appropriate misfit measure (here L_2 traveltime residual δt norm), the static kernel in (15) and the instantaneous sensitivity in (11) can be linked together.

Static traveltime kernel and waveform sensitivity are coupled through

$$K_d(\mathbf{x}) = \frac{1}{N_r} \int_0^T w_r(t) v(\mathbf{x}_r, t) \tilde{K}_d(\mathbf{x}, t) dt, \quad (16)$$

where the boxcar function $w_r(t)$ equals 1 inside the window where the sought phase arrives and $v(\mathbf{x}_r, t)$ is the velocity seismogram due to (13) recorded at station r . This is derived in equivalence to the cross-correlation traveltime perturbation as a function of δu in Dahlen *et al.* (2000). A scaling factor

$$N_r = \int_0^T \omega_r(t) v^2(\mathbf{x}_r, t) dt, \quad (17)$$

is introduced to rescale the kernel so that it does not depend on the source magnitude.

Finally, if the misfit function is computed with the L_2 norm, the optimization problem associated with traveltime tomography is defined as follows for the combination of all traveltime observations:

$$\text{find } \mathbf{m} \quad \text{such that} \quad \min \frac{1}{2} |t(\mathbf{m}) - t_{\text{obs}}|^2, \quad (18)$$

where $t(\mathbf{m})$ is solution of eq. (1). This expression allows eq. (15), once the problem is discretized on an inversion grid, to be solved in least-square sense as $\delta t_i = G_{ij} \delta d_j$, where the matrix G_{ij} contains the discrete sensitivity kernels through the integral (15).

3.4 Sensitivity kernel computational outline

Nissen-Meyer & Fournier (submitted) describe the work-flow to compute volumetric sensitivity kernels using the AXISEM software. Here, we extend their procedure to allow the computation of discontinuity kernels $K_d(x)$, introducing further numerical optimization described in Appendix B. By collapsing the equation of

motion onto a 2-D grid, AXISEM reduces the storage-space requirement so that it becomes possible to save the propagating wavefield in time at any resolution scale. An analogous process for similar resolution would not be feasible if a full 3-D solver was used.

The moment tensor calculation for eq. (1) in AXISEM is subdivided into four different simulations for as many independent moment tensor components M_m . Separate contributions, owing to the linearity of eq. (1), can be simply summed to recover the full moment-tensor wavefield (Nissen-Meyer *et al.* 2007b). The same property applies to kernels:

$$\mathbf{u} = \sum_m M_m \mathbf{u}_m \rightarrow K = \sum_m K_m \quad (19)$$

(valid for both boundary and volumetric kernels).

Due to the high-order accuracy in space (fourth order) and by honouring the CFL condition, space-and-time derivatives are computed at runtime as well as the other quantities in (11). Aware of the C_0 feature of the wavefield at the CMB, we checked the conservation of the normal component of the traction, $\mathbf{t} \cdot \hat{\mathbf{n}}$ across the solid–fluid boundary and the error of the second-order space derivatives. Derivatives are stable up to a certain simulation time, after that some discrepancies may rise. In general, P waves are always fine, and S waves travelling very long might be affected. The displacement gradient in the axi-symmetric domain is defined in eq. (A8). Its knowledge, together with that of the surface gradient and divergence (6), completely define the discontinuity kernel according to expression (11).

From the semi-disc-shaped domain (Fig. 11), we store, as a 1-D array, the values along the desired interface (a semi-circumference) in time. The simplest approach is to generate the spherical-kernel mesh, straight out of the spectral-element mesh. This latter procedure has the advantage of preserving the \hat{s} and \hat{z} components of AXISEM mesh points, acting only on the $\hat{\phi}$ component. Thus, no interpolation for any field variable is needed. One disadvantage is

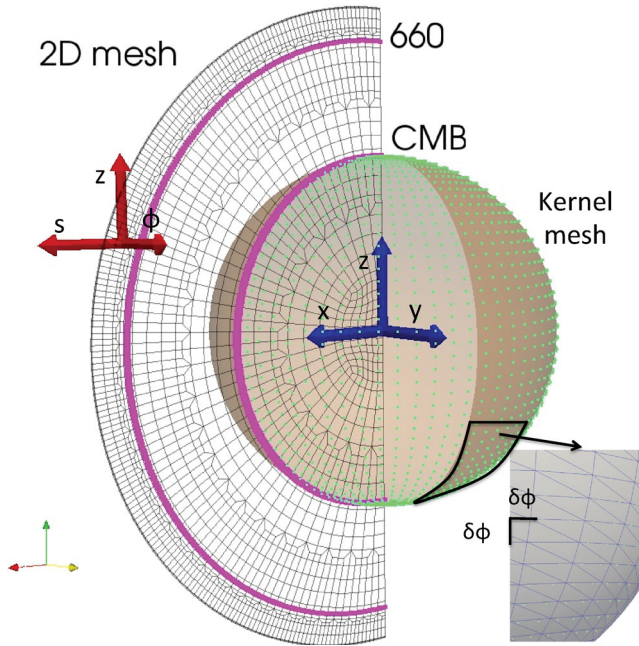


Figure 11. The kernel mesh (we removed a slice for clarity reasons) is constructed starting from the interface layer of the spectral element mesh (the gridded semi-disc). Here, for instance, the CMB layer is depicted. The spacing between $\Delta\phi$ is chosen to respect minimum point spacing. Magenta semi-circumferences represent the interface on the AXISEM mesh.

the point-clustering at the poles. Our procedure as well as the reference frame involved are illustrated in Fig. 11 and can be outlined as follows:

(i) The cylindrical coordinates of the point lying on the discontinuity (highlighted in magenta in Fig. 11) are copied from the AXISEM mesh of Fig. 11. By construction, they will all have $\hat{\phi}$ component equal to zero.

(ii) We select an appropriate $\Delta\phi$ and, moving in counter-clockwise direction, we loop over ϕ from 0 to 2π . The value of $\Delta\phi$ is such that the spatial resolution along the colatitude direction is equal to the longitudinal one at the equator (dedicated box in Fig. 11).

(iii) The grid-point coordinates are converted from cylindrical coordinates (\hat{s} , $\hat{\phi}$, \hat{z}) to Cartesian (\hat{x} , \hat{y} , \hat{z}) using eq. (A6).

For convenience, we define the kernel mesh with the z -axis pointing north, the same convention used in AXISEM. Both forward and backward fields are north-pole-oriented when saved from AXISEM, while they must be rotated when computing the kernel to their actual source or receiver positions. This rotation of the grid-point coordinates to the actual source location (ϕ_s , θ_s) or receiver location (ϕ_r , θ_r) boils down to coordinates and quantities to be re-mapped and will be described later. In our procedure, each grid point is associated with three different coordinate values: one for the kernel mesh, one for the forward field and one for backward. If we define the position of the new coordinate system as \mathbf{x}' (either for source or receiver), we may write: $\mathbf{x}' = \mathbf{R}\mathbf{x}$, where \mathbf{R} is defined for a given source or receiver position, (ϕ_s , θ_s) and (ϕ_r , θ_r) in eq. (A6).

The cylindrical frame is defined starting from the Cartesian frame through eq. (A7) such that a rotation of this latter produces the same rotation on the first. The coordinate basis vectors \hat{s} and \hat{z} correspond to the Cartesian \hat{x} and \hat{z} for $\phi = 0$ and $y = 0$, respectively. To each grid points corresponds a unique point onto the 2-D semi-disc as shown in Fig. A1. Vice-versa, to each 2-D mesh point correspond multiple kernel mesh-points along each azimuth.

The projection of the 3-D Cartesian point onto the 2-D semi-disc is achieved through eq. (A7). Hence, we run a linear search over the kernel mesh points to obtain the permutation array mapping the north-pole-oriented reference frame to the new desired position. The permutation always maps from the kernel mesh (3-D) to the AXISEM boundary mesh (2-D). The points from the 3-D mesh are projected onto the 2-D semi-disc. For each kernel grid-point (i -th), we measure the Euclidean distance from the 2-D points and a linear search to find the closer (j -th). The permutation array P for grid points i is: $P(i) = j$.

The azimuthal pre-factors applied to the 2-D wavefield serve to reconstruct the 3-D field and are computed using ϕ values. Fig. A2 depicts ϕ values for a given source-receiver pair (ϕ_s , θ_s) and (ϕ_r , θ_r). Eq. (A1) in Appendix A shows how pre-factors are applied to vector and tensor quantities depending on the moment tensor term currently computed. When the 3-D wavefield has been calculated, we can convert the field values to Cartesian coordinates. Again, the procedure uses the corresponding ϕ value for forward and backward fields. This is achieved applying a change of basis matrix $\mathbf{B}(\phi)$ as described in Appendix A. Wavefields now in Cartesian coordinates are written for two different coordinate systems: one with the z -axis pointing towards the source location and the other pointing towards the receiver.

Before calculating (15), we unify forward and backward field reference systems to the north-pole-oriented reference frame of the kernel mesh applying the same rotation matrix \mathbf{R} used to calculate

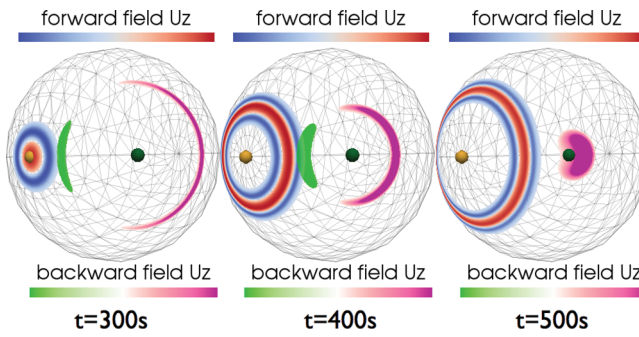


Figure 12. Forward and backward fields on the CMB surface for different moment in time. The fields are rotated to a Cartesian frame pointing north with z -axis. The forward field is centred at the north pole, whereas backward at $35^\circ\text{N } 0^\circ\text{E}$.

rotated coordinates.

$$\text{vectors: } \mathbf{u}_{\text{cart}} = \mathbf{R}^T \mathbf{u}_{\text{cyl}}, \quad \text{tensors: } \mathbf{E}_{\text{cart}} = \mathbf{R}^T \mathbf{E}_{\text{cyl}} \mathbf{R}. \quad (20)$$

The entries r_{ij} of \mathbf{R} are defined in Appendix A. Note that this rotation maps source-/receiver-rotated Cartesian frame to the north-pole-rotated frame, while the mesh coordinate's rotation previously defined mapped north pole oriented to source/receiver rotated. This difference boils down taking the transpose of the rotation matrix \mathbf{R} . For scalar quantities, no rotations are necessary as they are invariant. The results of these procedures are the forward and backward fields as if we were to use any other 3-D code and they are depicted for different time step in Fig. 12. Finally, the sensitivity kernel due to moment tensor M_m can be computed for any desired seismogram window.

4 BOUNDARY SENSITIVITY KERNELS

The software described in Appendix B provides boundary sensitivity kernels according to the procedure described in Section 3.4. We apply this to compute kernels at dominant periods of 10 and 25 s, using the source time function introduced in Section 2.1. For the sake of simplicity, we limit ourselves to monopole (explosive) sources, and accordingly only consider compressional waves. The sampling rate of modelled seismograms and wavefields by AXISEM is selected after a number of preliminary experiments. As noted by Fichtner *et al.* (2009), different sampling rates must be compared to identify the minimum rate threshold so that time aliasing of the seismic signal and of the sensitivity pattern is prevented. We typically sample the signal 10 times per dominant period, but we verify that this rate may be further reduced to limit the number of time steps needed to calculate the convolution in eq. (11).

The time integration window in eq. (16) is chosen such that the whole signal is enclosed inside, equivalent to Section 2.5 for the cross-correlation. We tuned specifically the size for every case to avoid the inclusion of other signals that may spoil the kernel-phase signature. A negative (positive) sensitivity value means that a positive topography anomaly will delay (accelerate) the arrival of the phase on the seismogram, and vice-versa, a depression would produce a positive (negative) arrival.

Most of our computations are carried out at the Swiss Center for Scientific Computing (<http://www.cscs.ch>). The computational cost for this not yet optimized algorithm (i.e. number of processors and wall-clock time) depends on the shortest modelled period, and on the depth range of the modelled discontinuity. It spans from few minutes on four processors for the CMB kernels at 25 s to

~ 1 hr on 24 processors for 10-s upper-mantle-discontinuity sensitivity kernels. Upper-mantle discontinuities are generally more expensive because of the different spatial sampling used to mesh the upper versus the lower mantle when designing the spectral element grid (Nissen-Meyer *et al.* 2008).

4.1 Transition-zone discontinuities

The seismic phases $P410P$, $P520P$ and $P660P$ (Fig. 13) associated with underside reflection of P waves at the ~ 410 , ~ 520 and ~ 660 km discontinuities, respectively, have been extensively used to map their topography (e.g. Gu *et al.* 1998; Flanagan & Shearer 1999; Gu & Dziewonski 2002; Deuss 2009). We determine the frequency-dependent sensitivity of these phases to discontinuity topography using an explosive event at a depth of 100 km. Because the source is isotropic, kernels are symmetric with respect to the vertical plane defined by source-receiver-centre of the Earth. The $P660P$ sensitivity kernel in Fig. 14 is characteristically X-shaped (Lawrence & Shearer 2008). Using differential traveltimes, the effects of structure elsewhere (crust and upper mantle for instance) may be removed, focusing the sensitivity only on the boundary. As demonstrated by Dahlen (2005), taking the differential travel-time exactly relates to computing the differential sensitivity kernels $K_{PP} - K_{P660P}$. Because of the small magnitude of the K_{PP} on the boundary, it amounts to $K_{PP} - K_{P660P} \simeq -K_{P660P}$ as Fig. 15 witnesses. The same results may be obtained using other underside reflection precursors such as $S660S$, $P410P$ or $S410S$. In particular, secondary-wave precursor is well suited for transition-zone studies because the impedance contrast for S waves is higher than for P waves with a consequently stronger signature in the seismogram providing better signal-to-noise ratio.

Our AXISEM-based kernels can be compared with those found by Dahlen (2005) and Lawrence & Shearer (2008) using the ray-theory finite-frequency approximation, and those found by Liu & Tromp (2008) using the adjoint technique combined with a 3-D spectral-element method. In spite of the difficulty of making direct comparisons when many parameters may differ slightly (source-time

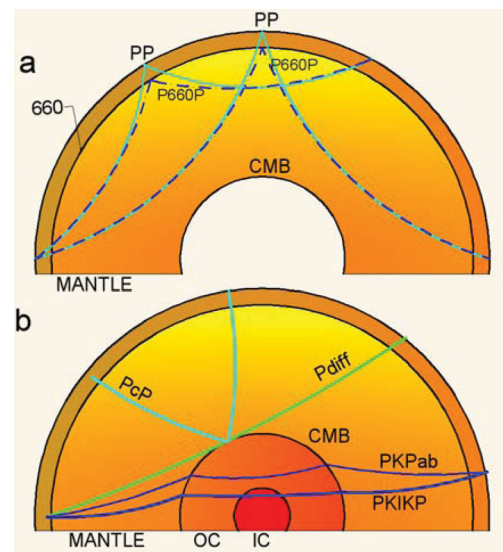


Figure 13. (Top) Cartoon showing the ray trajectories of the PP phase and its precursor discussed in Section 4.1. (Bottom) Ray trajectories of the core-reflected, core-refracted and core-diffracted phases of Section 4.2.

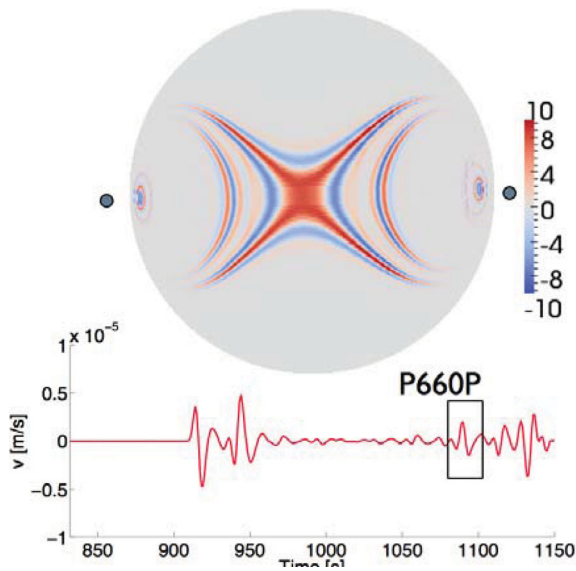


Figure 14. The K_{P660P} kernel for 140° epicentral distance and its associated vertical velocity component for dominant period of 10 s. Black dots mark source (depth 100 km) and receiver locations. Values are all scaled in $10^{-4} \text{ s km}^{-3}$. The kernel is projected on the ‘660’ surface.

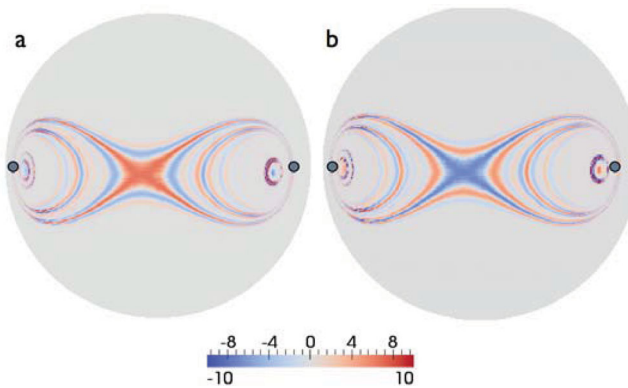


Figure 15. The K_{P660P} kernel for 95° epicentral distance on the left and the $K_{PP} - K_{P660P}$ on the right. The dominant period is 10 s. Black dots depict the source (depth 100 km) and receiver location. Values are all scaled in $10^{-4} \text{ s km}^{-3}$. The kernels are projected on the ‘660’ surface.

function, dominant period, source-receiver locations, etc.), our results are generally consistent with earlier findings (Fig. 19).

4.2 CMB discontinuity

The PcP , P_{diff} and various branches of PKP traveltimes (Fig. 13) have been extensively used to infer properties of lower mantle and CMB in many previous studies (Morelli & Dziewonski 1987; van der Hilst *et al.* 1998; Vasco *et al.* 1999; Boschi & Dziewonski 2000; Tanaka 2010; Soldati *et al.* 2012).

We calculate CMB sensitivity using an explosive event at a depth of 650 km. We mainly consider P waves and we record the vertical component of the seismograms. Again, since the source is isotropic, kernels are symmetric with respect to the vertical source-receiver plane. Results in Section 2.5 show the high sensitivity of PcP phases to the magnitude of the topography in the range between 30° and 70° epicentral distance. The sensitivity kernels displayed in Fig. 16

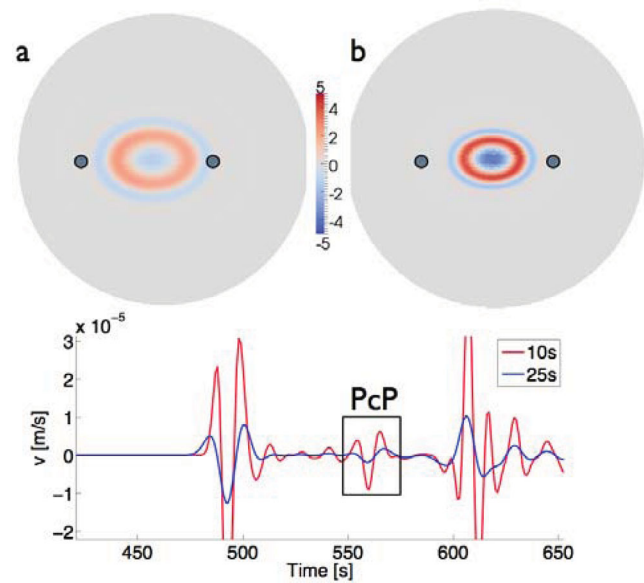


Figure 16. The PcP boundary sensitivity kernel for 25 s (left) and 10 s dominant periods (right) for an epicentral distance of 60° . It is elongated in the source-receiver direction, the area of non-negligible sensitivity extending up to 2000 km around the predicted bounce point for the 25 s case and roughly 1000 km for the 10 s case. The two dots indicate the location of source (depth 650 km) and receiver. Values are scaled to $10^{-5} \text{ s km}^{-3}$.

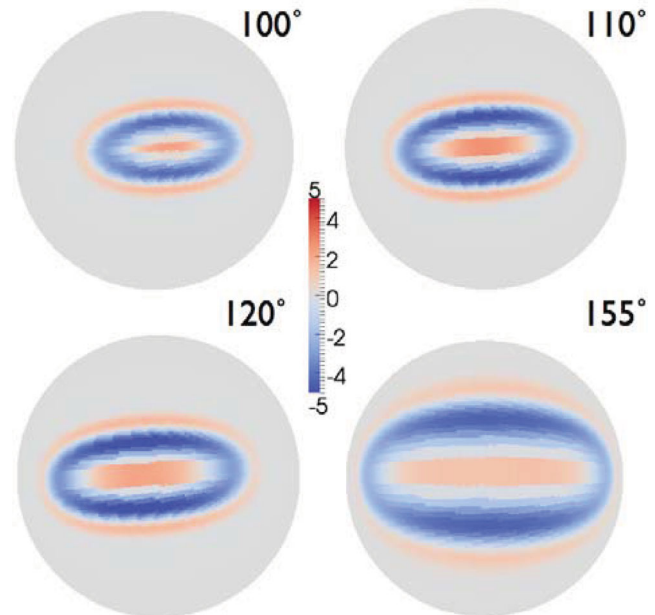


Figure 17. P_{diff} sensitivity kernels associated with different epicentral distances, as specified, at 25 s dominant period. Note how the sensitivity imprint becomes wider with growing epicentral distance. Values are scaled to $[10^{-5} \text{ s km}^{-3}]$. The depth of the event was 650 km.

point in the same direction as the synthetic test of Section 2.5, supporting that the bulk of the sensitivity is clustered in a relatively small area that slowly enlarges as the frequency decreases. We next present (Figs 17 and 18) finite-frequency sensitivity kernels associated with P_{diff} , for which (unlike, e.g. PcP) no asymptotic solution is available. The large width of the area where sensitivity

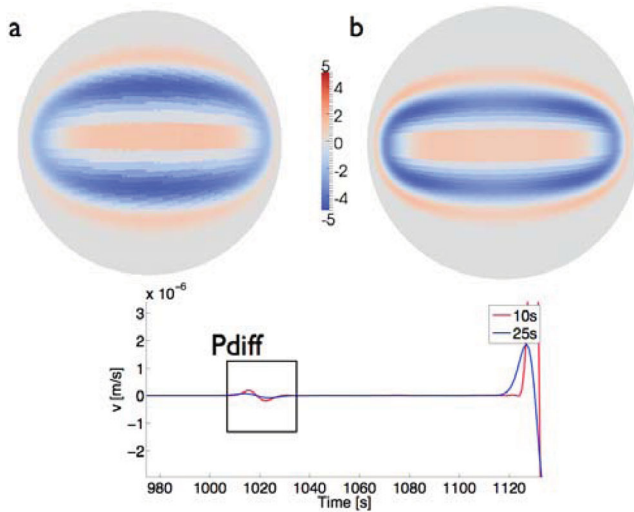


Figure 18. The P_{diff} sensitivity kernel for 155° epicentral distance (the maximum for P_{diff} being 157°), for (left) 25 s and (right) 10 s dominant period. The vertical component of our modelled trace is very weak at 25 s period, but it becomes clearer as the frequency increases. The same is true of this phase's sensitivity to CMB topography. Values are scaled to 10^{-5} s km $^{-3}$. The depth of the event was 650 km.

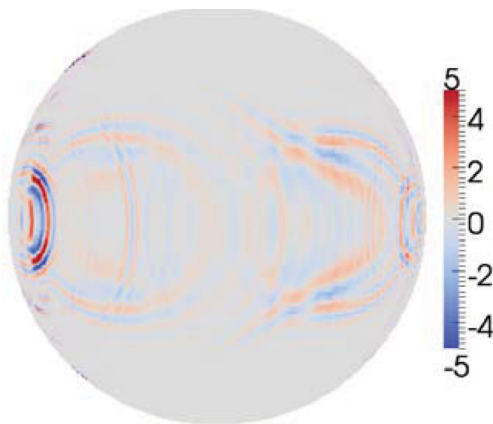


Figure 19. The figure shows the SKIKS kernel for an epicentral distance of 140° for 5 s dominant period. The entry point is located on the left side where the sensitivity is larger. A smaller sensitivity characterizes the exit point (on the right side). The pale signature on the background may be the effect of other SKS branches arriving almost simultaneously for this distance–depth range. Values are scaled to 10^{-4} s km $^{-3}$. The depth of the event was 650 km.

is non-negligible confirms our finding that this phase is sensitive to large-scale variations rather than small-scale (results found also in Section 2.5) and that positive topography might result in a delayed arrival time depending on its position relative to the source–receiver plane. Fig. 17 shows this phase for a suite of epicentral distances spanning its entire existence-range for 25 s dominant period. In spite of the seismic signal being extremely small for 155° , the P_{diff} sensitivity kernel can still be computed as shown in Fig. 18. Finally, the sensitivity of the PKIKP phase, and of the *ab* branch of PKP to CMB topography is illustrated in Fig. 20. Sensitivity is limited to the vicinity of the associated ray-path entry and exit points into and from the core. The pattern is similar, but the sign reversed with respect to P_{cP} sensitivity. Positive topography at and close to the entry/exit points causes a delayed arrival. As a final demonstration of the capability of our algorithm, we show in Fig. 21 the sensitivity kernel

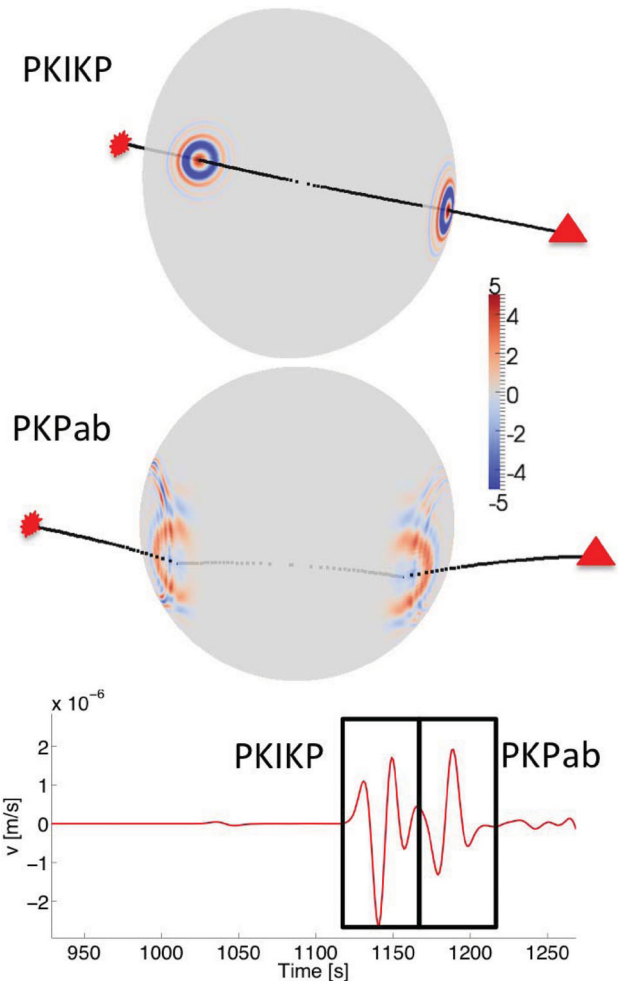


Figure 20. PKIKP and PKPab for 160° epicentral distance and a dominant period of 10 s. For the PKIKP, the plot is taken looking at a cross-section throughout the CMB to ease the visualization of both entry and exit points sensitivity. The black dotted line represents the theoretical ray path. The explosions and triangle marks, respectively, indicate a epicentre and receiver location. Values are scaled to 10^{-5} s km $^{-3}$. The depth of the event was 650 km.

for SKIKS for a dominant period of 5 s. Because of the epicentral distance chosen, other SKS branches arrive simultaneously. A $M_{r\phi}$ dipole moment-tensor source was used as source and the azimuthal pre-factors (A_1) were applied to the wavefield.

Comparison with other techniques both for P_{cP} and P_{diff} , although only qualitatively possible, may be done looking at the work from Dahlen (2005) and Liu & Tromp (2008). We also computed with SPECFEM3D GLOBE a P_{diff} kernel for 25 s dominant period (Fig. 19). The sensitivity pattern is similar to the one shown in Fig. 17. Small differences are attributed to the different algorithm used to solve eq. (1), the implementation of the adjoint source, the numerical time sampling, the selection of the time window and the meshing strategy.

4.3 Drawing connections: traveltime residual reconstructed with sensitivity kernels

One way to validate AXISEM-computed kernels and to show their applicability in tomography consists in reverse (Born) modelling

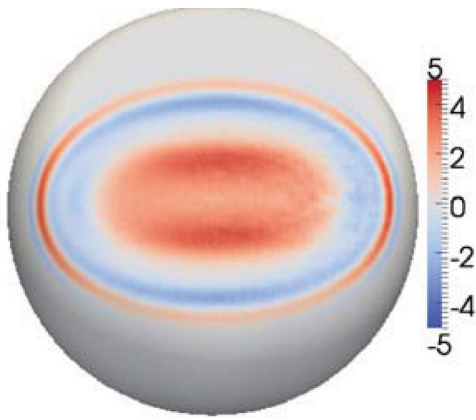


Figure 21. The P_{diff} sensitivity kernel at 155° calculated using the adjoint method implementation of SPEC-FEM3D GLOBE. The depth of the event was 650 km. Values are scaled to $10^{-5} \text{ s km}^{-3}$. This results may be compared with what obtained in Fig. 17.

the cross-correlated time residual using the kernels computed in the previous section. The first-order relationship (eq. 15), $\delta t = \int_{\Sigma} \delta d(x) K_d(\mathbf{x}) d\mathbf{x}^2$, is calculated evaluating the integral over a finite set of N constant basis functions (e.g. Boschi & Dziewonski 1999) $i \in \{1, \dots, N\}$ that reduces to a scalar product between two vectors:

$$\delta t = G_i \delta d_i. \quad (21)$$

Clearly, this holds for one specific event-topography-receiver configuration contained in the database causing the traveltime anomaly δt . Selecting one configuration of the many proposed throughout Section 2.5, we can compare reverse-modelled δt obtained through eq. (21) with previously computed SPEC-FEM3D-based quantities. We then evaluate the magnitude difference by constructing residual maps. For the sake of correctness, we ignore values $\leq 0.2 \text{ s}$ as they may be strongly influenced by small errors. The traveltime maps are depicted in Figs 22a and b. Fig. 22c shows that AXISEM-based reverse models slightly but systematically overestimates SPEC-FEM3D synthetics. The reason for this is presumably twofold: on the data acquisition side, the cross-correlation may underestimate such small signals, and the sampling rate (each 0.1 s) of the seismogram may not be sufficient. On the model side, the chain of approximations (including first order) used to build the discrete model parameter vector m_i , the gradient G_i and the small rounding used to construct the curved surface inside SPEC-FEM3D may play an important role. The consequences of this discrepancy are a slight underestimation of the magnitude of the topography at inversion time. We, however, infer that accuracy of AXISEM kernels is sufficient for application to typical (noisy and non-uniformly distributed) real seismic observations.

5 DISCUSSION

This paper addresses the effect of global seismic boundary topography on waveforms as well as sensitivity kernels, that is, the basis for tomographic imaging within the framework of full-wave theory. In the first part, we computed a large ground-truth database to examine the parameter space spanned by boundary perturbation geometry, epicentral distance, seismic phase, and frequency on seismograms. This is to be seen as a guide for optimal data configurations to illuminate such topography. In particular, PcP phases are, as expected, useful to detect CMB topography, whereas P_{diff} are less sensitive to

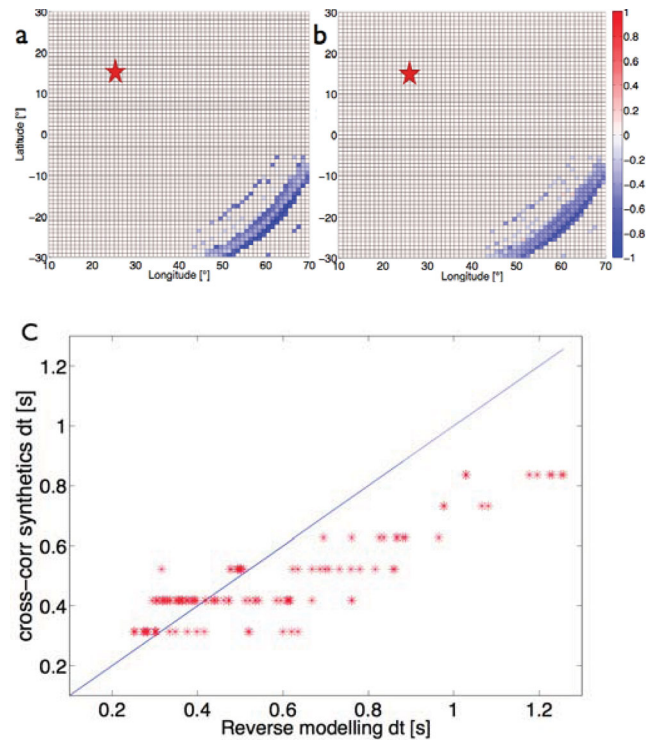


Figure 22. The results of the reverse modelling applying eq. (21) in panel (a) and those obtained by cross-correlating SPEC-FEM3D seismograms (b). While the patterns are practically identical, the magnitude of the δt differs as shown in the last plot (c). The continuous line represents perfect matching between the two measures, while asterisks the actual measured residuals. Cross-correlation residuals are divided in discrete band because of the finite sampling of the seismograms, while those from reverse modelling are more scattered around. We intentionally left out from plot (c) values lower than 0.2 s as they are not reliable.

such undulations. Our synthetic experiments highlight the importance of reflected phases to map discontinuity topography. The poor sensitivity shown by diffracted waves may represent two pitfalls of this analysis: $T_0 = 10 \text{ s}$ is not sufficient to constrain CMB topography (also for medium-scale structure) and traveltime analysis probably may not represent the ideal misfit measurement. A method based on phase envelopes as in Bozdog *et al.* (2011) might amplify the differences. Repeating this analysis with stronger topography will certainly lead to stronger signals but would violate the constraint that the majority of CMB-sensitive observations impose on their magnitude. The database has been computed using full 3-D wave propagation, and as such the prohibitive cost at high resolution (e.g. below periods of 10 s) prevents us from an analysis covering the full spectrum. As tomographic inversions must be conducted at sufficiently high frequency (below 10 s) to detect this topography, we must rely on computationally more efficient methods than full 3-D wave propagation to construct sensitivity kernels.

In the second part, we implemented boundary sensitivity kernels using first-order perturbation theory following Dahlen (2005), using the axisymmetric spectral-element code AXISEM. This methodology allows to model frequencies as high as required by the database analysis, even in the framework of large-scale tomographic data sets. Those kernels, representing the Fréchet derivatives of perturbations of data with respect to topography, account for finite-frequency effects and can be used in waveform inversion to invert for boundary perturbation. Our kernels share similar properties with

kernels calculated in other studies (Hung *et al.* 2000; Dahlen 2005; Peter *et al.* 2007; Liu & Tromp 2008). The dependency of the area of non-negligible sensitivity and amplitude upon the period is evident in all cases considered here. Its extension scales with the wavelength and it asymptotically collapses to the ray size at infinite frequency. The values of maximum and minimum sensitivity change correspondingly to accommodate the area change. Varying the epicentral distance produces similar effects. From the perspective of Section 4, in our future finite-frequency imaging of the CMB, the inclusion of PKP should also be considered as the sensitivity footprint (particularly true for PKIKP) is similar to PcP. The reverse modelling illustrated in Section 4.3 confirms the robustness of our approach against more general 3-D wave propagation.

The inclusion of volumetric perturbation combined with topography will be the next step towards waveform inversion accounting for the strong connection between boundary topography and the surrounding volumetric structure. Mantle and crust corrections are known to heavily trade-off with topography in the target area as discussed by Flanagan & Shearer (1999). Combining the boundary sensitivity algorithm with the work from Nissen-Meyer & Fournier (submitted) will allow a joint volumetric-boundary full-waveform inversion work-flow at global scale, scalable to any resolution. Before reaching this stage, further time has to be devoted to assessing the resolution power of this methodology based on spherical Earth models. This will be achieved by inverting the synthetic seismograms calculated in Section 2.5 and verifying how well input structures are recovered. A study on waveforms as in Section 2.5 can be extended to the combined effect of boundary plus volumetric anomalies: first to assess how the two trade off one another and later to benchmark the joint waveform inversion.

Before implementing the inversion technique, further numerical and computational optimization is needed. In spite of the efficiency of the software by Nissen-Meyer *et al.* (2007b), computing global sensitivity kernels at periods smaller than 5 s is quite expensive. Preliminary application to GPU hardware shows that GPU computing may be a key factor to making this endeavour affordable, besides reverting to frequency-domain convolutions (Nissen-Meyer & Fournier submitted).

ACKNOWLEDGMENTS

We thank reviewers Paul Cupillard and Robert Geller for their careful reviews as well as the editor's comments that helped to improve the manuscript. Computations were undertaken on the Brutus Cluster at ETH Zurich and at the Swiss National Supercomputing Center (CSCS). We thank both centres for their support. We acknowledge funding from QUEST, an international training network by the European Commission, and Petaquake, a project of the Swiss high-performance, high-productivity platform HP2C.

REFERENCES

Bijwaard, H., Spakman, W. & Engdahl, E.R., 1998. Closing the gap between regional and global travel time tomography, *J. geophys. Res.*, **103**(B12), 30 055–30 078.

Becker, T.W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models. *G-cubed*, **3**, 2001GC00068.

Boschi, L. & Dziewonski, A., 1999. High and low resolution images of the earth's mantle—implications of different approaches to tomographic modeling, *J. geophys. Res.*, **104**, 567–594.

Boschi, L. & Dziewonski, A., 2000. Whole earth tomography from delay times of P, PcP and PKP phases: lateral heterogeneities in the outer core or radial anisotropy in the mantle?, *J. geophys. Res.*, **105**, 13 675–13 696.

Bozdag, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**, 845–870.

Dahlen, F.A., 2004. Resolution limit of traveltime tomography, *Geophys. J. Int.*, **157**, 315–331.

Dahlen, F.A., 2005. Finite-frequency sensitivity kernels for boundary topography perturbations, *Geophys. J. Int.*, **162**, 525–540.

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

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

Della Mora, S., Boschi, L., Tackley, P.J., Nakagawa, T. & Giardini, D., 2011. Low seismic resolution cannot explain S/P decorrelation in the lower mantle, *Geophys. Res. Lett.*, **38**, L12303, doi:10.1029/2011GL047559, 2011.

Deuss, A., 2009. Global observations of mantle discontinuities using SS and PP precursors, *Surv. Geophys.*, **4**(30), 301–326.

Dziewonski, A. & Anderson, D.L., 1981. Preliminary Reference Earth Model, *Phys. Earth planet. Inter.*, **25**, 297–356.

Dziewonski, A. & Gilbert, J.F., 1976. The effect of small, aspherical perturbations on travel time and a re-examination of the corrections for ellipticity, *Geophys. J. R. astr. Soc.*, **235**(44), 7–17.

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**(3), 1703–1725.

Flanagan, M.P. & Shearer, P.M., 1999. A map of topography on the 410-km discontinuity from pp precursors, *Geophys. Res. Lett.*, **26**, 549–552.

Forté, A.M., Mitrovica, J.X. & Woodward, R.L., 1995. Seismic-geodynamic determination of the origin of excess ellipticity of the core-mantle boundary, *Earth planet. Sci. Lett.*, **9**(22), 1013–1016.

Geller, R.J. & Hara, T., 1993. Two efficient algorithms for iterative linearized inversion of seismic waveform data, *Geophys. J. Int.*, **115**, 699–710, doi:10.1111/j.1365-246X.1993.tb01488.x.

Gu, Y. & Dziewonski, A., 2002. Global variability of transition zone thickness, *J. geophys. Res.*, **107**(B7), doi:10.1029/2001JB000489.

Gu, Y., Dziewonski, A. & Agee, C.B., 1998. Global de correlation of the topography of transition zone discontinuities, *Earth planet. Sci. Lett.*, **157**(1–2), 57–67, doi:10.1016/S0012-821X(98)00027-2.

Gutenberg, B., 1914. Über erdbebenwellen. viia. Beobachtungen an Registrierungen von Fernbeben in Goettingen und Folgerungen über die Konstitution des Erdkörpers, *Nachr. Ges. Wiss. Goettingen, Math-Phys. Klasse*, **7**, 1–52 and 125–176.

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

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

Ita, J.J. & Stixrude, L., 1992. Petrology, elasticity and composition of the transition zone, *J. geophys. Res.*, **97**, 6849–6866.

Jackson, A., Bloxham, J. & Gubbins, D., 1993. Time-dependent flow at the core surface and conservation of angular momentum in the coupled core-mantle system, in *Dynamics of Earth's Deep Interior and Earth Rotation*, pp. 97–107, eds LeMouél, J.-L., Smylie, D.E. & Herring, T., number 72 in AGU Geophysical Monograph, IUGG, Washington, DC.

Koelemeijer, P.J., Deuss, A. & Trampert, J., 2012. Normal mode sensitivity to earth's D'' layer and topography on the core-mantle boundary: what we can and cannot see, *Geophys. J. Int.*, 553–568, doi:10.1111/j.1365-246X.2012.05499.x.

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, **149**, 390–412.

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. 3-D models, oceans, rotation, and self-gravitation, *Geophys. J. Int.*, **150**, 303–318.

Lamb, H., 1904. On the propagation of tremors over the surface of an elastic solid, *Phil. Trans. R. Soc. Lond.*, **203**, 1–42.

Lawrence, J.F. & Shearer, P.M., 2008. Imaging mantle transition zone thickness with SdS-SS finite-frequency sensitivity kernels, *Geophys. J. Int.*, **174**, 143–158.

- Lehmann, I., 1936. Inner earth, *Bur. Cent. seism. Int.*, **14**, 3–31.
- Li, C., Van der Hilst, R.D., Engdahl, E.R. & Burdick, S., 2008. A new global model for P wave speed variations in Earth's mantle, *Geochem. Geophys. Geosyst.*, **9**, doi:10.1029/2007GC001806.
- Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.*, **174**, 265–286.
- Montelli, R., Nolet, G., Dahlen, F.A. & Masters, G., 2006. A catalogue of deep mantle plumes: new results from finite-frequency tomography, *Geochem. Geophys. Geosyst.*, **7**, Q11007, doi:10.1029/2006GC001248.
- Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R. & Hung, S.-H., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343, doi:10.1126/science.1092485.
- Mora, P., 1989. Inversion=migration + tomography, *Geophysics*, **54**(12), 1575–1586.
- Morelli, A. & Dziewonski, A., 1987. Topography of the core-mantle boundary and lateral homogeneity of the liquid core, *Nature*, **325**, 678–683.
- Nissen-Meyer, T., Dahlen, F.A. & Fournier, A., 2007a. Spherical-earth Fréchet sensitivity kernels, *Geophys. J. Int.*, **168**, 1051–1066.
- Nissen-Meyer, T., Fournier, A. & Dahlen, F.A., 2007b. A 2-D spectral-element method for computing spherical-earth seismograms—I. Moment-tensor source, *Geophys. J. Int.*, **168**, 1067–1093.
- Nissen-Meyer, T. & Fournier, A., Time-frequency sensitivity for global tomography, *Geophys. J. Int.*, submitted.
- Nissen-Meyer, T., Fournier, A. & Dahlen, F.A., 2008. A 2-D spectral-element method for computing spherical-earth seismograms—II. Background models, *Geophys. J. Int.*, **174**, 873–888.
- Obayashi, M., Yoshimitsu, J. & Fukao, Y., 2009. Tearing of stagnant slab, *Science*, **324**(5931), 1173–1175.
- Oldham, R.D., 1906. The constitution of the earth, *Q. J. geol. Soc. Lond.*, **62**, 456–475.
- Peter, D., Tape, C., Boschi, L. & Woodhouse, J.H., 2007. Surface wave tomography: global membrane waves and adjoint methods, *Geophys. J. Int.*, **171**, 1098–1117.
- Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2010. S40RTS: a degree-40 shear velocity model for the mantle from new rayleigh wave dispersion, teleseismic traveltimes, and normal-mode splitting function measurements, *Geophys. J. Int.*, **184**, 1223–1236, doi:10.1111/j.1365-246X.2010.04884.x.
- Simmons, N.A., Forte, A.M., Boschi, L. & Grand, S.P., 2010. GyPSuM: a joint tomographic model of mantle density and seismic wave speeds, *J. geophys. Res.*, **115**, B12310, doi:10.1029/2010JB007631.
- Soldati, G., Boschi, L. & Forte, A., 2012. Tomography of core-mantle boundary and lowermost mantle coupled by geodynamics, *Geophys. J. Int.*, **189**, 730–746, doi:10.1111/j.1365-246X.2012.05413.x.
- Soldati, G., Boschi, L. & Piersanti, A., 2003. Outer core density heterogeneity and the discrepancy between PKP and PcP travel time observations, *Geophys. Res. Lett.*, **30**, doi:10.1029/2002GL016647.
- Takeuchi, N., 2005. Finite boundary perturbation theory for the elastic equation of motion, *Geophys. J. Int.*, **160**, 1044–1058.
- Tanaka, S., 2010. Constraints on the core-mantle boundary topography from P4KPPcP differential travel times, *J. geophys. Res.*, **115**, 57–71.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern california crust, *Science*, **325**, 988–992.
- Tarantola, A., 2005. *Inverse Problem Theory*, Society for Industrial and Applied Mathematics, Philadelphia.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
- van der Hilst, R.H., Widiyantoro, S., Creager, K.C. & McSweeney, T.J., 1998. Deep subduction and aspherical variations in P-wavespeed at the base of the earth's mantle, *Geodyn. Ser.*, **28**, 5–20.
- Vasco, D.W., Johnson, L.R. & Marques, O., 1999. Global Earth structure: inference and assessment, *Geophys. J. Int.*, **137**(2), 381–407.

APPENDIX A: AXIAL SYMMETRY AND CARTESIAN REFERENCE FRAME

Going from the 2-D semi-disc quantities given by AXISEM to general 3-D cylindrical coordinates requires some multiplicative factor to be applied:

$$\text{vectors } \mathbf{u}_{\text{cyl}}^{3D}(s, \phi, z) = (f_1(\phi)u_s^{2D}, f_2(\phi)u_\phi^{2D}, f_1(\phi)u_z^{2D}), \quad (\text{A1})$$

$$\text{tensor } \mathbf{E}_{\text{cyl}}^{3D}(s, \phi, z) = \mathcal{F} \otimes \mathbf{E}^{2D}, \quad (\text{A2})$$

where \mathcal{F} contains the azimuthal pre-factors f_1 and f_2 :

$$\mathcal{F} = \begin{pmatrix} f_1(\phi) & f_2(\phi) & f_1(\phi) \\ f_2(\phi) & f_1(\phi) & f_2(\phi) \\ f_1(\phi) & f_2(\phi) & f_1(\phi) \end{pmatrix}, \quad (\text{A3})$$

and \otimes stands for element-wise product. The values of the azimuthal pre-factors f_i depend on ϕ , source type and moment tensor component. The following table summarizes the dependencies.

Table 1.

Source type	M component	Prefactor	
		$f_1(\phi)$	$f_2(\phi)$
Monopole	$\frac{1}{2}M_{zz}$	1	0
Dipole	$\frac{1}{2}(M_{xx} + M_{yy})$	1	0
	$\frac{1}{2}M_{xz}, p_x$	$\cos \phi$	$-\sin \phi$
Quadrupole	$\frac{1}{2}M_{yz}, p_y$	$\sin \phi$	$\cos \phi$
	$\frac{1}{2}(M_{xz} - M_{yy})$	$\cos 2\phi$	$-\sin 2\phi$
	M_{xy}	$\sin 2\phi$	$\cos 2\phi$

The angle ϕ is known for each of the kernel grid points \mathbf{x} , projecting its components onto the x - y plane and calculating the angle with the *atan2* function. The Cartesian reference frame is assumed to have the z -axis oriented towards the source (Fig. 11). The values of ϕ have to be recalculated for every source-receiver configuration. The easier case happens when all sources are monopole and therefore azimuth-invariant such that this operation can be avoided. If we denote with \mathbf{B} the matrix that maps the Cartesian reference frame into the cylindrical reference frame, then the following transformations hold:

$$\text{vectors: } \mathbf{u}_{\text{cart}} = \mathbf{B}^T \mathbf{u}_{\text{cyl}}^{3D} \quad \text{tensors: } \mathbf{E}_{\text{cart}} = \mathbf{B}^T \mathbf{E}_{\text{cyl}}^{3D} \mathbf{B}, \quad (\text{A4})$$

where the b_{ij} entries of the change of basis matrix \mathbf{B} are:

$$\mathbf{B} = \begin{pmatrix} \cos \phi_r & -\sin \phi_r & 0 \\ \sin \phi_r & \cos \phi_r & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{A5})$$

For example, the cylindrical coordinates of the kernel mesh are converted to Cartesian with: $\mathbf{x}_{\text{cart}} = \mathbf{B}^T \mathbf{x}_{\text{cyl}}^{3D}$. The coefficients for the rotation matrix \mathbf{R} in eq. (20) are:

$$\mathbf{R} = \begin{pmatrix} \cos \theta_{r,s} \cos \phi_{r,s} & -\sin \phi_{r,s} & \sin \theta_{r,s} \cos \phi_{r,s} \\ \cos \theta_{r,s} \sin \phi_{r,s} & \cos \phi_{r,s} & \sin \theta_{r,s} \sin \phi_{r,s} \\ -\sin \theta_{r,s} & 0 & 0 \end{pmatrix}. \quad (\text{A6})$$

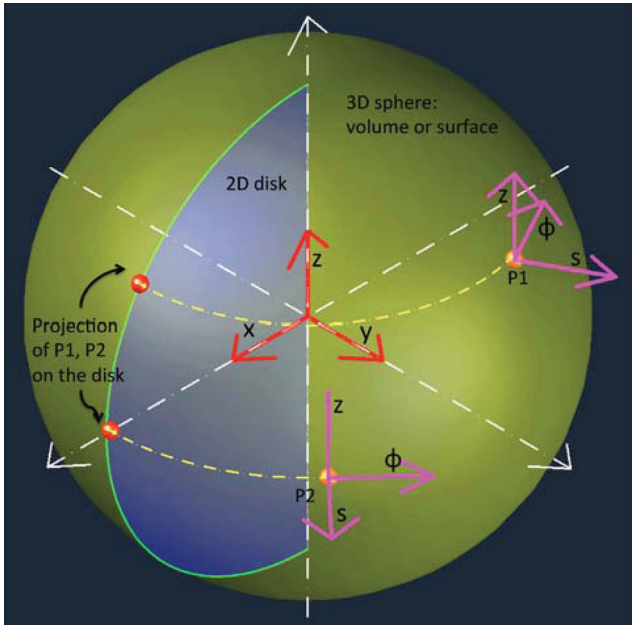


Figure A1. A visual interpretation of the projection defined in (A7). The point on the semi-disc is chosen to be the closest to the actual projection.

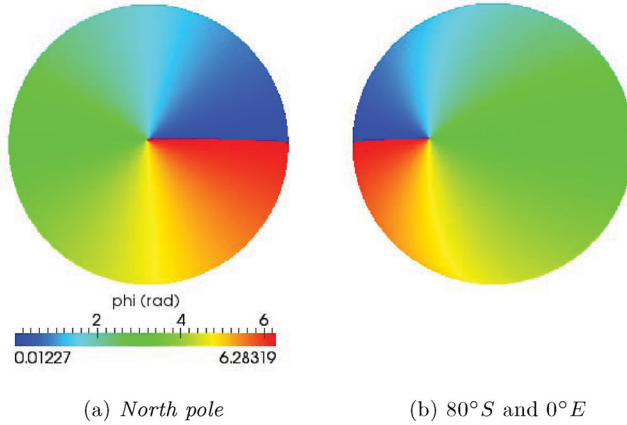


Figure A2. The kernel mesh with ϕ values superimposed. (Left) Values of ϕ for the forward field located at north pole. (Right) The backward field rotated 170° towards south. The z -axis in both plots points outwards from the page.

Here, \mathbf{R} is uniquely defined for all the mesh grid points \mathbf{x} depending on the source and receiver coordinate (θ_s, ϕ_s) or (θ_r, ϕ_r) . Note that (θ_r, ϕ_r) are not the initial receiver coordinates, but rather those defined within the coordinate frame attached to the kernel mesh.

Finally, the geometrical relationships between Cartesian 3-D grid (x, y, z) and 2-D semi-disc in cylindrical coordinates (s, ϕ, z) are:

$$s = \sqrt{x^2 + y^2}; \quad (\text{A7a})$$

$$\phi = \tan^{-1}(y/x); \quad (\text{A7b})$$

$$z = z. \quad (\text{A7c})$$

We compute the gradient of the displacement on-the-fly inside AXISEM via

$$\nabla \mathbf{u} = \begin{bmatrix} \frac{\partial u_s}{\partial s} & \frac{1}{s} \frac{\partial u_s}{\partial \phi} & \frac{\partial u_s}{\partial z} \\ \frac{\partial u_\phi}{\partial s} - \frac{u_\phi}{s} & \frac{1}{s} \frac{\partial u_\phi}{\partial \phi} + \frac{u_s}{s} & \frac{\partial u_\phi}{\partial z} \\ \frac{\partial u_z}{\partial s} & \frac{1}{s} \frac{\partial u_z}{\partial \phi} & \frac{\partial u_z}{\partial z} \end{bmatrix}, \quad (\text{A8})$$

keeping always in mind that multiplicative pre-factors have to be applied once the geometry of the kernel mesh is known.

APPENDIX B: ALGORITHMIC OPTIMIZATION

To compute K_d in (15) for a given database including N_r receivers and N_s sources, we proceed as follows:

- Forward fields for every source depth s are computed via AXISEM, saving the necessary wavefields 10 times per dominant period of the source time function \hat{m} , that is, re-sampling using a larger time step, then imposed by the numerical CFL condition. This value ensures sufficient time sampling to avoid both aliasing and reduce the computational cost and storage size (Fichtner *et al.* 2009).
- By virtue of axial symmetry and the seismic representation theorem, only two simulations for all the observations are sufficient to compute the backward field.
- The kernel code computes eq. (11) according to Section 3.2 for every given receiver r associated to each source s .
- Results are stored and the procedure is repeated for all the sources in the database. Eventually, once one has the inversion grid available, inversion matrix entries may be given as output.

The parallelization of the forward solver AXISEM is discussed in Nissen-Meyer *et al.* (2007b, 2008); therefore, we do not discuss it here, but depict the domain decomposition in Fig. B1.

B1 Parallelization and in-core optimization

In our formulation, computing kernels is a task involving wavefield rotations and convolution in time domain; hence, the parallelization is straightforward. We set up the domain decomposition along

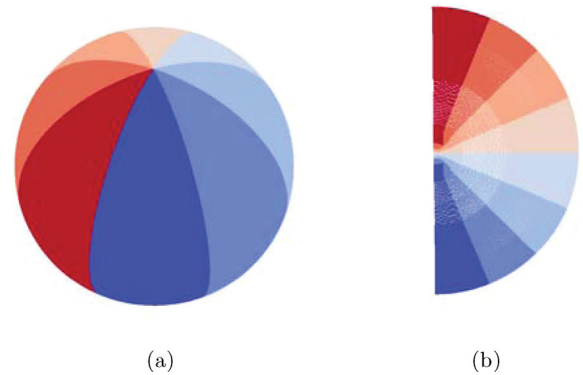


Figure B1. Domain decomposition for the kernel mesh (a) and for AXISEM spectral element mesh (b) for eight processors as an example. Each slice is handled by a different processor. The number of processor is defined by the users according to memory/wall-clock time desired.

the ϕ direction assigning each processor a vertical slice spanning a given interval. The decomposition ensures perfect load balance for optimal scalability on large parallel machines. To calculate only the kernel, there is no need for intra-node communication so that convolution, once the domain is split, is handled completely independently by each processor: Every processor simply needs to know the $\delta\phi$ spanned by the assigned slice and the permutation array to remap for actual source or receiver location. Remaining in the same framework, one can also think to directly project the kernel onto the inversion grid. Significant performance gain was obtained by embedding a GPU-accelerated region that handle this task with extremely high efficiency. This has proven to be particularly useful to calculate G_i in eq. (21). ‘In-core’ optimizations were obtained designing the convolution as 1-D array operations compliant to the Fortran-like memory layout. Wherever possible we avoid nested for-cycles or control statements. The code is designed to exploit the most recent compiler optimization especially those of PGI with GPU accelerators (www.pgroup.com).

B2 I/O optimization

For every source depth s , AXISEM saves vector fields that are used to compute kernels. For I/O performance, we mean all the operations transferring data back and forth from the central memory (RAM)

to the hard disk. A few best practices have been followed when designing this I/O intensive code for parallel machines.

- Use single precision whenever double is not necessary (almost always true in numerical wave propagation). This countermeasure simply halves storage requirements.
- I/O infrastructures are designed to handle large files. I/O of small files is never optimized; thus, it is better to reduce the number of total disk access (disk latency time is several orders higher than RAM access). Maximize bandwidth usage for disk I/O building memory buffers inside RAM memory. To compute one boundary kernel with AXISEM, only a 1-D time-dependent collection of points is required (aligned along the interface) and it is therefore possible to fit all the necessary field values inside a node memory. In such a way, input fields are read from disk once-and-for-all at the start-up.
- Use a portable format that can be written/read from heterogeneous system. For example, data from the forward simulation can be computed on specifically designed machines and then the output fields need to be read from another machine that may have a different hardware structure to specifically enhance the performance of the algorithm calculating kernels. We use the NetCDF data library (www.ucar.org), a self describing, meta-data format widely supported and optimized for data-intensive application.
- We embed data visualization in the output data format using the CF convention (<http://cf-pcmdi.llnl.gov/>), natively supported by software like Paraview or GMT.