Geophysical Journal International

Geophys. J. Int. (2014) **198,** 55–71
Advance Access publication 2014 April 29
GJI Seismology

doi: 10.1093/gji/ggu112

Seismic waveform inversion for core–mantle boundary topography

Andrea Colombi,[1,*] Tarje Nissen-Meyer,[1,†] Lapo Boschi[2,3] and Domenico Giardini[4]

[1] *Institut für geophysik, ETH Zürich (CH), Switzerland. E-mail: colombia@ethz.ch*
[2] *UPMC, Universit Pierre et Marie Curie, ISTEP, Institut des Sciences de la Terre de Paris, F-75005 Paris, France*
[3] *CNRS, Centre National de la Recherche Scientifique, UMR 7193, F-75005 Paris, France*
[4] *Institut für Geophysik, NO H 69.1 Sonneggstrasse, CH-58092 Zürich (CH), Switzerland*

Accepted 2014 March 20. Received 2014 March 19; in original form 2013 December 4

SUMMARY

The topography of the core–mantle boundary (CMB) is directly linked to the dynamics of both the mantle and the outer core, although it is poorly constrained and understood. Recent studies have produced topography models with mutual agreement up to degree 2. A broad-band waveform inversion strategy is introduced and applied here, with relatively low computational cost and based on a first-order Born approximation. Its performance is validated using synthetic waveforms calculated in theoretical earth models that include different topography patterns with varying lateral wavelengths, from 600 to 2500 km, and magnitudes (∼10 km peak-to-peak). The source–receiver geometry focuses mainly on the $P_{\rm diff}$, *PKP*, *PcP* and *ScS* phases. The results show that *PKP* branches, *PcP* and *ScS* generally perform well and in a similar fashion, while $P_{\rm diff}$ yields unsatisfactory results. We investigate also how 3-D mantle correction influences the output models, and find that despite the disturbance introduced, the models recovered do not appear to be biased, provided that the 3-D model is correct. Using cross-correlated traveltimes, we derive new topography models from both *P* and *S* waves. The static corrections used to remove the mantle effect are likely to affect the inversion, compromising the agreement between models derived from *P* and *S* data. By modelling traveltime residuals starting from sensitivity kernels, we show how the simultaneous use of volumetric and boundary kernels can reduce the bias coming from mantle structures. The joint inversion approach should be the only reliable method to invert for CMB topography using absolute cross-correlation traveltimes.

Key words: Inverse theory; Body waves; Seismic tomography; Wave scattering and diffraction; Wave propagation.

1 INTRODUCTION

The core–mantle boundary (CMB), is the strongest discontinuity in the earth interior, and it separates the solid mantle from the fluid outer core. Its topography is very likely related to both the thermal/compositional/viscosity structure and the associated convection of the mantle (Forte *et al.* 1995; Soldati *et al.* 2012, 2013), and also to the properties of the outer core, where vigorous convection is believed to generate the earth magnetic dynamo (e.g. Jackson *et al.* 1993). A number of studies, starting with Morelli & Dziewonski (1987), have mapped the CMB topography based on compressional wave traveltimes, while others (e.g. Li *et al.* 1991; Ishii & Tromp 2001) have inverted observations of eigenfrequency splitting, to focus on normal modes that are sensitive to the CMB. Most ray-based models feature a similar degree-2 pattern with peak-to-peak topography of a few kilometres, although a shorter wavelength structure is more difficult to constrain, as for volumetric tomography. Several studies have indicated a discrepancy in the CMB structure as mapped by core-refracted (various branches of the *PKP* phase) 'versus' core-reflected seismic waves (e.g. Boschi & Dziewonski 1999; Vasco *et al.* 1999; Soldati *et al.* 2003), which has cast some doubt on the validity of CMB maps derived from these data. Soldati *et al.* (2012), however, showed that the discrepancy is significantly reduced if the inverse problem solution for CMB topography is required to be coupled with the seismic structure in the mantle according to the theory of Forte *et al.* (1995).

The seismically slower fluid outer core and the complex 3-D structures above this region are the cause of triplication, caustics, scattering and attenuation, and these prevent the obtaining of good quality global seismic data; that is, with sufficient signal-to-noise ratio (Valenzuela & Wysession 1998). The few broadband observations available are normally clustered in particular regions,

*Now at: ISTerre, Université Joseph Fourier, Grenoble, BP 53 38041 Grenoble, CEDEX 9, France.
†Now at: Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 4RP, United Kingdom.

as, for instance, in the studies of Thorne *et al.* (2007) and Vanacore *et al.* (2010). From the installation of dense acquisition grids, such as the USArray (e.g. Garnero 2006), we can have access to a large amount of data for global/regional deep-mantle studies. These data can be used in conjunction with full numerical wave propagation algorithms (e.g. Peter *et al.* 2011) to feed gradient inversion techniques, with the help of the adjoint method (e.g. Tromp *et al.* 2005; Fichtner *et al.* 2008) for the calculation of frequency-dependent sensitivity kernels. For spherically symmetric background models, the work of Nissen-Meyer *et al.* (2008) provides a spectral element code that can be used to solve the wave equation down to a 5–10 s period with reasonable computational resources. This allowed us to compute the boundary sensitivity kernels for the 1-D background models in an efficient manner (Colombi *et al.* 2012), compared to other more expensive techniques for 3-D models (e.g. Liu & Tromp 2008; Fichtner *et al.* 2009). So far, only Lawrence & Shearer (2008) has constrained the upper-mantle discontinuities, following a similar, although ray theory-based, approach. One possible way to discriminate which among the various CMB-sensitive phases are best to constrain the topography is through sensitivity studies of the CMB phases using synthetic data calculated with large spectral element simulations through various topographic CMB models (Colombi *et al.* 2012). Core reflections (*PcP*, *ScS*) and core diffraction (*PKP*) provide better images of the CMB topography, while diffracted waves perform poorly. In contrast to differential traveltime measurements, like *PmKP* − *PcP* from Tanaka (2010), absolute traveltimes suffer the disturbing effects of the crust and mantle. The appropriate correction of these effects is key to the success, and this is a central challenge for such studies. Static corrections calculated through ray tracing are the most common method, and many databases (e.g. Ritsema *et al.* 2011) have been released along with the residual traveltime corrections for the crust and mantle structures. Although potentially capable of solving this issue, the simultaneous joint inversion (e.g. Vasco *et al.* 1995) has never been used in CMB studies. As it requires good data coverage, it has been applied more frequently in exploration seismology (Hobro *et al.* 2003), although without accounting for finite-frequency sensitivity.

In this paper, we implement the volumetric and boundary sensitivity kernels first introduced in Colombi *et al.* (2012) in an inversion scheme that relies on spherically symmetric background models. In Section 2, we discuss the implications of CMB topography on wave propagation, and we present various synthetic data sets for the *P* and *S* phases, used for testing and for benchmarks. Section 3 deals with the formulation of the matrix inverse problem for a single class of parameters (topography anomaly) or for the simultaneous joint inversion of the volumetric and boundary anomalies. Section 4 contains a suite of synthetic inversions for the variation of the topography, dominant period, parametrization and seismic phase. We also address the problem connected to the simultaneous effect of 3-D mantle plus topography when using an automated measuring technique for cross-correlation, and we analyse the effects of the source–receiver distribution. In Section 5, we make use of our algorithm to invert an actual data set that contains various phases, and we discuss the analogies and similarities with other studies, and the geodynamic implications. Finally, we use volumetric sensitivity kernels to try to discriminate between the mantle and the boundary contributions, and we argue that a joint inversion can resolve this trade-off. To complete this study, an Appendix details our approach to the solution of the inverse problem.

2 GLOBAL SYNTHETIC WAVEFORM DATABASE

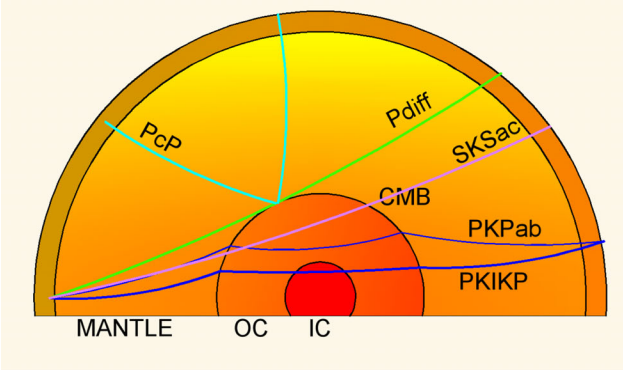
2.1 Global wave propagation

The Earth is a heterogeneous body that is characterized by strong discontinuities that cause significant complexities in the wavefield. At the global scale, the phenomenon of seismic wave propagation is well explained by linear elastodynamics theory. The presence of a fluid outer core is modelled by coupling the acoustic wave equation with the elastic wave equation through transmission conditions at the CMB (Komatitsch & Tromp 2002a; Nissen-Meyer *et al.* 2007b). These interface conditions are crucial for the derivation of the boundary sensitivity kernels with the approach of Dahlen (2005), using Born theory (Dahlen & Tromp 1998). At first order, spherical symmetric background models, such as the Preliminary Reference Earth Model (PREM; Dziewonski & Anderson 1981), contain all of the major discontinuities, and depth-dependent mechanical properties are sufficiently accurate for the computation of sensitivity kernels.

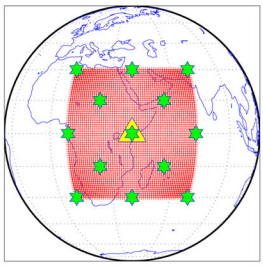
The effects of the earth rotation, gravitation and of the oceans, are all negligible at the seismic period (20–10 s) and the geographic scale length (20°–180°) considered in this study (Dahlen & Tromp 1998; Komatitsch & Tromp 2002b). Effects of the earth ellipticity are accounted for by applying a linear correction to the data, as shown by Dziewonski & Gilbert (1976). The seismic event is introduced in the wave equation in the form of a source–time function that is applied to the moment tensor $\mathbf{M}$, which contains information about the radiation pattern of the rupture. To highlight the structural effects of this methodology, we only consider explosive events for *P* waves, or dipoles (e.g. dip-slip) in the case of *S* waves.

2.2 CMB topography

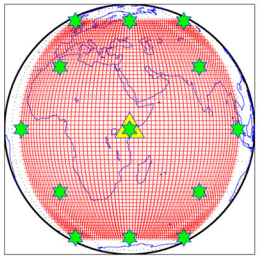
Most published models of CMB topography share the same degree-2 pattern, as in Soldati *et al.* (2012) and Tanaka (2010). The structures are in the order of a few thousands of kilometres, while smaller features are difficult to constrain, for several reasons: lack of coverage, noise and ray theory flaws. The sensitivity of seismic waves scales as a function of their frequency. Higher frequencies should be more sensitive to smaller topography (with respect to the lateral length scale). We determine whether a medium-to-small ($\sim$600 $\div$ 1000 km) topography at the global scale can be constrained successfully with 20 or 10 s dominant periods that amount to a length scale over the wavelength (hereafter abbreviated with γ) ratio of $\sim$3 $\div$ 10. If this is the case, it means that current models can be retrieved with much longer period signals, or vice versa, that with shorter period data and good coverage, existent CMB maps may be refined. We use two different anomaly patterns to determine whether positive/negative topography (Fig. 1f) introduces some bias with respect to the easiest case (Fig. 1e), where only positive topography was inserted. The magnitude, which is chosen according to previous estimates (see Koelemeijer *et al.* 2012, for a review of the models), is between 5 and 10 km from peak to peak. Figs 1(e) and (f) show that the topographies inserted into the model lie within the small perturbation regime. Smaller topography values could be potentially retrieved as long as we can measure data with a sufficient signal-to-noise ratio and periods $\ll$10 s. This is only true as long as one can efficiently account for the effect of the overlying 3-D structure that for such a small value of topography may introduce a strong bias as it is shown in Section 6.



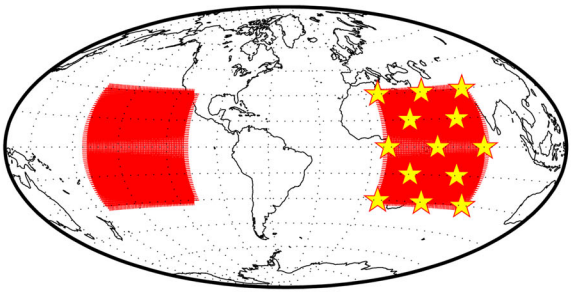
(a) *Ray-path*



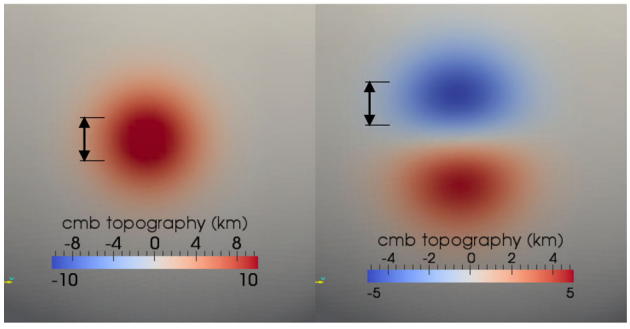
(b) PcP, ScS



(c) P_{diff}



(d) $PcP + PKP$



(e) *Bell-shaped uplift* (f) *Depression/topography*

Figure 1. (a) Ray path for the CMB seismic phases considered in our database. (b–d) Source–receiver layout for, respectively, the *PcP*, $P_{\rm diff}$ and $PKP + PcP$ phases. The red dots show the stations, and the stars represent the events. The spacing between each station is 1° in both directions for *PKP* and *PcP*, and 2° for $P_{\rm diff}$. (e and f) Two CMB map views: a bell-shaped topography is obtained by projecting a Gaussian distribution, and its first derivative, along one major arc. For each spectral element simulation, we modify the lateral extension and magnitude of the topography. A sample of the scale used to define the extension is plotted next to the topography.

Table 1. The number of usable traveltime observations ($N\Delta t(\cdot)$) for each of the inversions upon the topography size and dominant period. *PKP* represents *PKIKP* + *PKPbc*.

Phase topography	Model	$N\,\Delta t$(10 s)	$N\,\Delta t$(20 s)
PcP ∼1000 km × 10 km	PREM	3561	2691
PcP ∼600 km × ±5 km	PREM	3321	2411
ScS ∼1000 km × 10 km	PREM	4385	3657
ScS ∼600 km × ±5 km	PREM	4185	3375
$P_{\rm diff}$ ∼1000 km × 10 km	PREM	2988	None
$P_{\rm diff}$ ∼2500 km × ±5 km	PREM	2718	None
PKP ∼1000 km × 10 km	PREM	2532	1825
PKP ∼600 km × ±5 km	PREM	2347	1762
PcP ∼1000 km × 10 km	3-D	9450	7459
PcP ∼600 km × ±5 km	3-D	8610	6753

2.3 Selected seismic phases and acquisition grid

Our synthetic database contains several data sets (Table 1), and is an extension of that introduced in Colombi *et al.* (2012). The software SPECFEM3D GLOBE (Komatitsch & Tromp 2002a) was used to calculate the perturbed and unperturbed waveforms with an accuracy down to a period of 10 s. The perturbation consists of an anomaly δr (positive or negative) of different magnitudes, shapes and lateral length scales superimposed on the native global mesh (Figs 1e and f) at the CMB. The mantle is approximated either with the PREM (Dziewonski & Anderson 1981) or with a 3-D model (Ritsema *et al.* 2011).

Traveltime residuals are calculated with cross-correlations, applying the procedure described in Colombi *et al.* (2012). The output of this operation is a traveltime residual map $\Delta t(\theta, \phi)$ for each seismic phase and earthquake, which is pass-band filtered at the desired dominant period (in this study, 10 or 20 s). The cross-correlation is applied on a time window that is centred at the predicted arrival of the phase considered, and calculated using TauP-2.0 (http://www.seis.sc.edu/software/TauP/). For explosive sources, we consider only the vertical components of the seismogram, while for those generating *S* waves, we look at the transverse component (*ScS*).

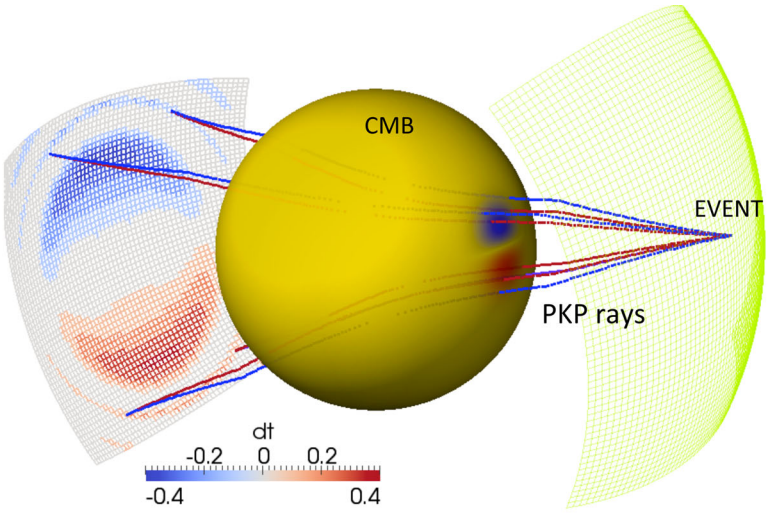
The database used in Colombi *et al.* (2012) contains only *PcP* and $P_{\rm diff}$, while here, to broadly cover the spectrum of energy partitions at the CMB, we also calculated the *PKP* branches and the *ScS*. The ray paths of the phases used are shown in Fig. 1(a).

The *PKP* branches and *PcP* have been used in several studies to constrain the CMB topography (Morelli & Dziewonski 1987; Soldati *et al.* 2003, 2012; Sze & Van der Hilst 2003), or the small-scale structure of the lower mantle (Lay & Garnero 2004; Vanacore *et al.* 2010) in local or regional studies. Tanaka (2010) considered differential traveltimes from $PmKP - PcP$ to eliminate the mantle effects. As these parameters require more than double the CPU time, they will not be considered here. The use of $P_{\rm diff}$ is not that common, because modelling $P_{\rm diff}$ correctly using ray theory requires the measure of the onset time, and hence high frequencies (e.g. Richards 1973), which are not present in diffracted waves (e.g. Valenzuela & Wysession 1998). Global data from the *ScS* phases are available from the study of Ritsema *et al.* (2011), although these data had yet to be inverted for CMB topography. *S*-wave reflections appear to be easily identifiable at longer periods (10 s), because of the higher reflection coefficient at the CMB than the *PcP*. These phases have very different propagation paths, as seen in Fig. 1(a). To capture them correctly, we arranged the source–receiver layout in different ways upon the corresponding epicentral distance.

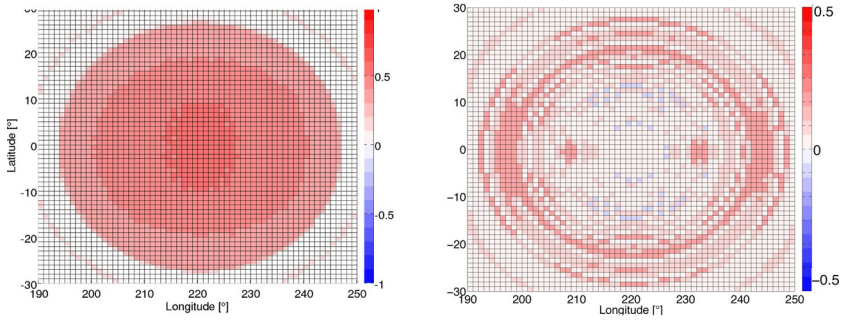
Colombi *et al.* (2012) carried out an extensive broad-band analysis for multiple frequencies over the *PcP* and P_{diff} data (see, for instance, figs 8–10 of Colombi *et al.* 2012). While for the *ScS* data the sensitivity is very similar to the *PcP* (provided that we consider the same wavelength and hence different periods), interesting results are obtained from *PKP* branches. Especially for *PKIKP*, the sensitivity kernels look very similar to those of *PcP* and *ScS*, even if they are double-sided due to their different entry and exit points. The residuals for one of the source–receiver configurations are shown in Figs 2(b) and (c). These are measured from the vertical components for the *PKIKP* and *PKPbc* phases that were filtered at the 10 s dominant period. If infinite frequencies were used, the magnitude of the residual traveltime would be approximately half of that expected for *PcP* (e.g. Morelli & Dziewonski 1987), because the sensitivity is partitioned into two sides (Colombi *et al.* 2012), for an equal anomaly δr. The reason why only *PKIKP* carries a strong signature in our experiment is twofold: as Fig. 2(a) shows, the *PKPbc* ray (blue) passes slightly off the topography because the take-off angle is not as high as for *PKIKP* (red), which are sensing the anomaly closer. Secondly, the sensitivity kernels show milder sensitivities compared to *PKIKP*, which results in a smoothed effect. The residuals for 20 s (not shown here) lead to similar considerations, even if they are smaller in magnitude because of the frequency-dependent sensitivity. It is reasonable to expect a similar behaviour when dealing with *SKS* phases, which are not considered in the synthetic database for computational reasons, although they are used for the actual inversion.

2.4 Background models

Knowledge of the mantle structure is a pre-requisite to any CMB topography study. The bulk of the database was calculated using PREM as the background model, to highlight the topographic effects first. To give a more realistic imprint for the *PcP* data, we also considered a recent 3-D mantle model. SPECFEM3D GLOBE embeds the *S*- and *P*-wave models by Ritsema *et al.* (2011) (S40RTS+P12), and hence it is feasible to simulate the effects of the full 3-D mantle and CMB topography altogether (Fig. 3). The *P* velocity profile is obtained by applying a scaling a depth-dependent factor $\mathrm{d}v_s/\mathrm{d}v_p$ between 1.25 and 4 (Ritsema & Van Heijst 2002) with respect to S40RTS. Details on how these models are discretized can be found in Komatitsch & Tromp (2002b). The topography patterns are the same as those in Figs 1(a) and (b), paired with the *PcP* source–receiver layout. The question whether discontinuity topography can



(a) *Experimental layout*



(b) $PKIKP\ \Delta t$ (c) $PKPbc\ \Delta t$

Figure 2. (a) Ray trajectory for the *PKIKP* and *PKPbc* leave the epicentre towards the stations with a different take-off angle. The map of the CMB shows the topography (Fig. 1f) located beneath the source. The colour map on the receivers indicates the magnitude of the Δt. (b and c) Residual traveltimes for *PKIKP* and *PKPbc*, measured on the vertical component for the configuration in Fig. 1(e). The residuals are calculated with respect to a CMB with no topography. The dominant period is 10 s, and the event was located at a depth of 100 km, on top of the topography.

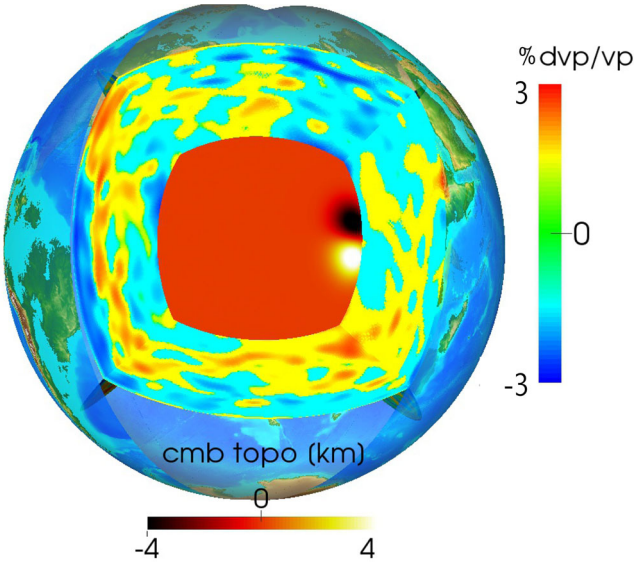


Figure 3. The v_p structure of the S40RTS+P12 model discretized to resolve waves down to 10 s with the topography inserted at the CMB, after removing one of the cubed sphere chunks characterizing the SPECFEM3D GLOBE mesh. This configuration is used to reproduce the joint effects of the topography and the 3-D mantle.

be mapped also in the presence of realistic mantle perturbations can be addressed in a straightforward way. For each source–topography–receiver configuration, we run one simulation with PREM, another with only the 3-D mantle, and another with the 3-D mantle plus the topography, while the pure effect of the topography was calculated in Section 2.3. If the mantle has little influence on the measurements, the effects of the topography remain more or less the same. We used cross-correlation to measure the traveltime residual $\Delta t^{\text{topo}}_{\text{PREM}}$ for *PcP*. We explained in Section 2.3 how $\Delta t^{\text{topo}}_{\text{PREM}}$ is measured, and we add the subscript PREM here as a reminder that a 1-D background model was used. Here we extend this procedure to:

(i) $\Delta t^{\text{notopo}}_{\text{3-D}}$ is the traveltime residual due solely to S40RTS, with respect to the PREM;

(ii) $\Delta t^{\text{topo}}_{\text{3-D}}$ is the joint effect of the CMB topography plus the mantle traveltime residual, with respect to the PREM.

If the mantle does not affect the measurement of the $\Delta t^{\text{topo}}_{\text{PREM}}$, the following relationship holds:

$$\Delta t^{\text{topo}}_{\text{PREM}} \simeq \Delta t^{\text{topo}}_{\text{3-D}} - \Delta t^{\text{notopo}}_{\text{3-D}}. \quad (1)$$

Fig. 4 shows the result of this analysis for the *PcP* case. At first sight, the effects of the 3-D mantle structure have a large effect on Δt_{topo}, although considering that the magnitude of $\Delta t^{\text{notopo}}_{\text{3-D}}$ is up to 10-fold greater than $\Delta t^{\text{topo}}_{\text{PREM}}$ (from the ratio measured while exploring the database), the results obtained are good overall. The topography produces a maximum residual predicted by ray theory, of $\sim$0.15 s km^{-1} for the *PcP* phase. Because of the limited extension of the topography and the frequency-dependent sensitivity, this value is considered as a theoretical threshold hence our measurements will be smaller. Δt_{topo} in Fig. 4(b) is obtained by applying eq. (1). The difference between this latter and the reference $\Delta t^{\text{topo}}_{\text{PREM}}$ (Fig. 4a) is shown in Fig. 4(c). The strongest component of the signature is correctly retrieved (white stripe in the middle), while smaller magnitude residuals appear around the darker area. This might be considered as part of the error, either due to the automated measuring strategy or coupled to the effects of the topography and the 3-D mantle. In Section 4, we show that this small error does

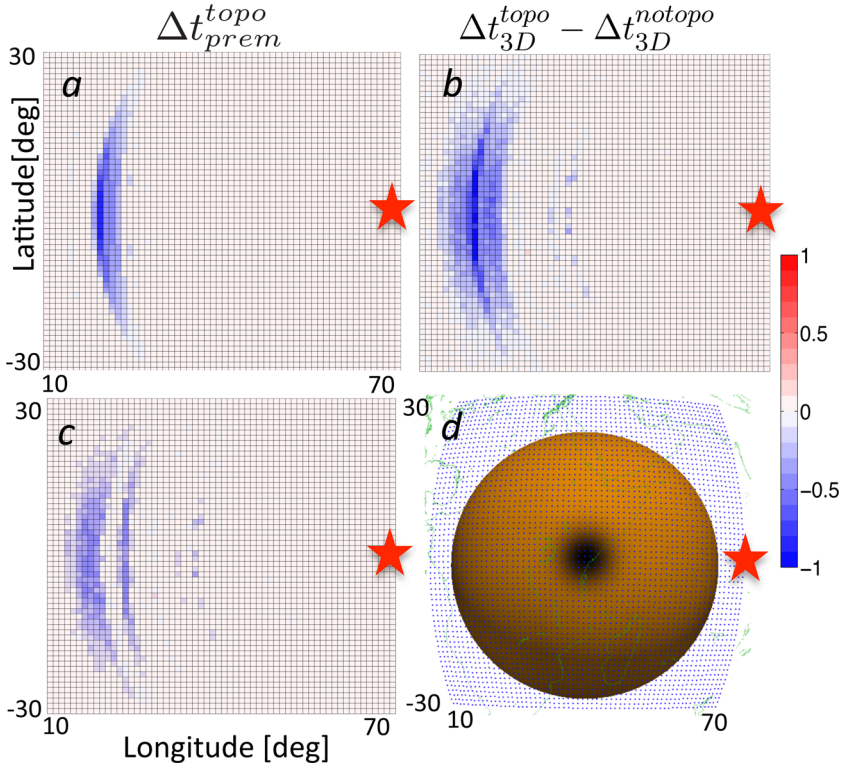


Figure 4. (a) The reference Δt calculated with the 1-D mantle model. (b) The traveltime residual maps induced by the CMB topography calculated applying eq. (1), using the synthetics computed for a 3-D mantle. (c) The difference between the reference in (a) and the 3-D-mantle-influenced residuals in (b). (d) The topography was located beneath the centre of the grid and the event (red star) at (0°N, 70°E).

not compromise the results of the inversions. During the tests at different frequencies, we observed an increasing number of outliers (i.e. residuals with unfeasibly high magnitudes) for the smaller dominant period. This was probably due to the occurrence of cycle skipping (Virieux & Operto 2009), an anomalous phase difference between reference and perturbed seismograms that induces cross-correlations to fail. We manually excluded these from the analysis in Fig. 4. The threshold was set the same as the dominant period of the data considered (in this case $\mathrm{d}t < 10$ s).

3 THE INVERSION ALGORITHM

We use the approach developed in Colombi *et al.* (2012) to compute the sensitivity kernels for each source–receiver pair used for the inversion. The inverse problem is then solved with a standard sparse matrix inversion. The procedure remains identical for synthetics and actual data inversions.

3.1 Waveform boundary sensitivity kernels

The central point of the study of Colombi *et al.* (2012) was the implementation of an efficient method to compute sensitivity kernels based on 1-D reference models on the Born approximation, such that the traveltime sensitivity to the boundary perturbation is defined by the following integral:

$$\Delta t = \int_{\Sigma} \delta r(x)\, K_d(\mathbf{x})\, \mathrm{d}\mathbf{x}^2. \tag{2}$$

Sensitivity kernels for each source–receiver pair are then expanded on a set of $i = 1, \ldots, N$ base functions φ_i. Choosing 'pixel' parametrization as described by Boschi & Dziewonski (1999), with the local base functions that take unitary values only in a pixel, reduces eq. (2) to a simple scalar product between two vectors (making use of the Einstein summation convention):

$$\Delta t = G_i\, \delta r_i. \tag{3}$$

Clearly, this holds for one specific event–topography–receiver configuration contained in the database associated with the traveltime anomaly Δt. The difference between ray theory and finite-frequency tomography lies in the construction of G_i [see Tape *et al.* (2006) or Peter *et al.* (2007) for further details on these analogies]. The integral that appears in eq. (2) is solved by a quadrature with the trapezoidal rule (Quarteroni *et al.* 2007), which after discretization, amounts to a simple element-wise product:

$$G_i = K_i\, A_i, \tag{4}$$

where A_i is the ith pixel area. This approximation holds as long as the pixel size is small enough to sufficiently replicate the kernel pattern (see Fig. 6). If our database contains $j = 1, \ldots, M$ traveltime observations that we collect in a residual traveltime vector $\Delta\mathbf{t}$, the linearized inverse problem takes the following (matrix) form:

$$\Delta\mathbf{t} = \mathbf{G} \cdot \delta\mathbf{r}, \tag{5}$$

where $\mathbf{G}$ is an $M \times N$ matrix, and each row G_i corresponds to one discrete sensitivity kernel (3). The model update vector $\delta\mathbf{r}$ can be computed easily when $\mathbf{G}$ is not too large, using a least-squares method, which in its canonical form for mixed-determined problems takes the following form (Boschi & Dziewonski 1999):

$$\Delta\mathbf{r} = \left(\mathbf{G}^T \cdot \mathbf{G} + \alpha^2\mathbf{I}\right)^{-1} \cdot \mathbf{G}^T \Delta\mathbf{t}, \tag{6}$$

where the regularization term α helps to stabilize the inversion. The correct value of this term is chosen by exploring the L-curves (Fig. 5c). The sparse system is then solved with a conjugated gradient algorithm (Golub & Loan 1996). To further increase the smoothness of the solution, it is also possible to damp the roughness of the solution using an operator $\mathbf{B}$, which when applied to $\mathbf{m}$ gives the gradient of the misfit function (Boschi & Dziewonski 1999), and the final form of (6) is:

$$\delta\mathbf{r} = \left(\mathbf{G}^T \cdot \mathbf{G} + \alpha_1^2\mathbf{I} + \alpha_2^2\mathbf{B} \cdot \mathbf{B}^T\right)^{-1} \cdot \mathbf{G}^T \Delta\mathbf{t}, \tag{7}$$

where α_2 is the amount of roughness damping introduced in the inversion. Our tomographic inversion is linearized, which is an approximation that is assumed by most global tomographic models. Regardless of the method used to invert the matrix $\mathbf{G}$, as will be seen by our inversions, this is plausible for the small topographic perturbations on the CMB. In this study, we do not consider *a priori* information, data correlation matrices or *a posteriori* probabilities. The Fréchet derivative of the misfit function projected on the residual vector associated with the traveltime tomography is calculated as:

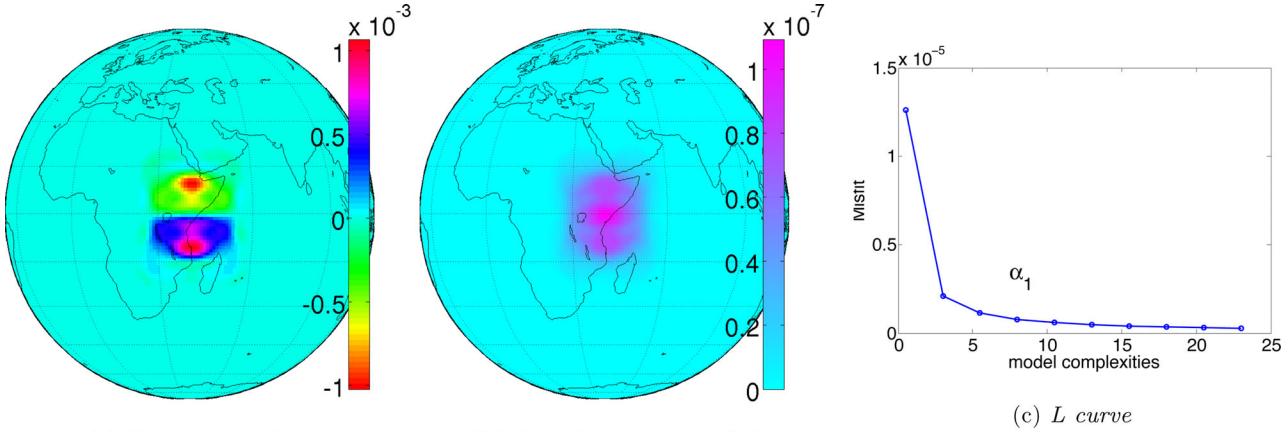
$$\mathbf{g} = -\mathbf{G}^T \Delta\mathbf{t}, \tag{8}$$

which gives information about the direction and the magnitude of the model updates. The L_1 norm of the columns of $\mathbf{G}$ gives the value of the hit count and is typically used as proxy for the data coverage, such as ray plot in the case of ray-based tomography. The L_1 norm in Fig. 5(b) can be compared with the hit count in Fig. 11(e), to appreciate the more uniform coverage provided by finite frequency. Because the residuals used for the inversions cover only a small area of the CMB, as the diagonal of the coverage indicates, the addition of the damping term α_1 makes the system solvable also in the case α_2 is set to 0 (i.e. this indicates that unconstrained pixels have no variations, hence $\delta r = 0$), although in most of the synthetic inversions $N \gg M$.

3.2 Continuous versus discrete kernels

The term 'continuous' here indicates that the mesh on which the kernel is calculated is much finer than the inversion mesh. The transition from continuous to discrete is illustrated in Fig. 6, where a sensitivity function $K(\mathbf{x})$ for the *PcP* phase is projected onto the inversion grid that is made of equal-area pixels. Fig. 6 highlights a key point of the inversion strategy: if the size of the inversion grid is not sufficiently small, the topology of the kernel is lost, as it is shown in Fig. 6(d). As demonstrated by Panning *et al.* (2009), using volumetric kernels, the size of the first Fresnel zone is directly related to the anomaly length scale we want to map. Along with the data coverage, this factor drives the choice of the inversion grid. We investigate the implication of these parameters in Section 4. For the synthetic inversion, we use various inversion grid parametrizations: 2°-, 3°-, 5°-sized pixels. Depending on the size of the first Fresnel zone, different parameterizations can have different effects on the results. As a general rule, we state that 2° and 3° work fairly well for *PcP*, *ScS* and *PKP*, and 5° is insufficient. As the P_{diff} sensitivity is larger, this allows us to use up to 5° large pixels. The number of model parameters given by each parametrization and the phases suited for each pixel size are summarized in Table 2.

It is not only the number of parameters that determines the time to solve the system (5), but also the size of the non-negligible sensitivity that dominates the sparsity of $\mathbf{G}$. The number of observations depends on the seismic phase, the topography and the dominant period used to filter the seismogram. For the same input topography, shorter periods are more sensible and, therefore, lead to higher



(a) *Projected gradient* (b) L_1 *columns norm of* $\mathbf{G}$ (c) *L curve*

Figure 5. Various features of the inversion strategy for the *PcP* data set. (a) Gradient of the misfit function to be minimized with the inversion. The gradient, as eq. (8), contains information about the magnitude and direction of the model updates. (b) L_1 norm of the columns of $\mathbf{G}$ which is used to explore the data coverage. (c) The L-curve employed to trim the damping parameter α_1 that is used to stabilize the inversion of the system in (6).

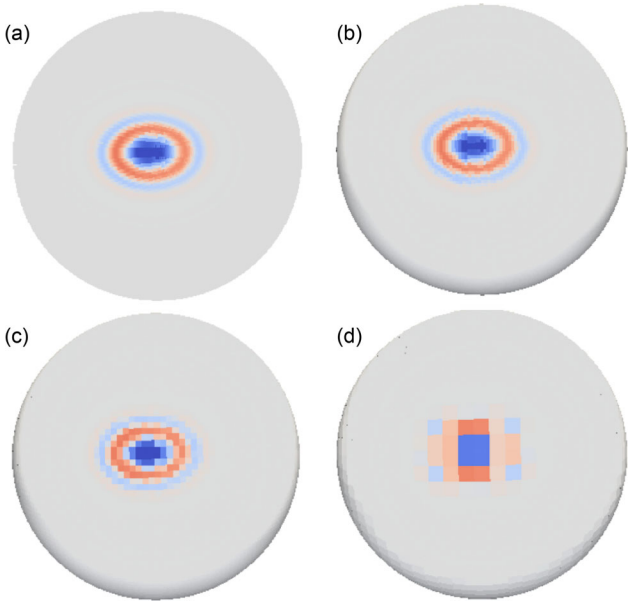


Figure 6. (a) Sensitivity kernel for the *PcP* phase at 50° epicentral distance for an explosive event with a 20-s dominant period. (b–d) The same kernels expanded on inversion grids with 2°, 3° and 5° large pixels, respectively. The pixel are constructed on an equal-area basis on a spherical surface.

Table 2. Depending on the size of the sensitivity area, different pixel sizes might be suited. We summarize here the parametrization that is best suited to our inversions.

Grid size	Model parameters	Phases suited
2°	10 316	ScS, PcP, P_{diff}, PKP, SKS
3°	4592	ScS, PcP, P_{diff}, PKP, SKS
5°	1656	P_{diff}

number of observations. Table 1 summarizes the number of synthetics that are inverted in the next section.

3.3 Volumetric and boundary sensitivity: the joint approach

A seismic waveform anomaly δu is associated with volumetric δv and boundary δr perturbations, through the corresponding volumetric and boundary sensitivity kernels $\tilde{K}_v$ and $\tilde{K}_d$:

$$\delta u(t) = \int_\Omega \delta v\, \tilde{K}_v(\mathbf{x}, t)\, \mathrm{d}\mathbf{x}^3 + \int_\Sigma \delta r\, \tilde{K}_d(\mathbf{x}, t)\, \mathrm{d}\mathbf{x}^2. \tag{9}$$

In this section, we focus on the volumetric term and couple its discrete form with the boundary sensitivity that was previously computed. Assuming an isotropic elastic earth, and fulfilling the approximation introduced in Colombi *et al.* (2012), we can relate the waveform perturbations for volumetric heterogeneities to material properties using: density ρ, Young's modulus λ and shear modulus μ. The volumetric sensitivity kernels calculated for the frequency domain by Nissen-Meyer *et al.* (2007a) can be computed for the time domain, and they can be written as follows:

$$\tilde{K}_\rho(\mathbf{x}, t) = \partial_t \vec{\mathbf{u}} * \partial_t \overleftarrow{\mathbf{u}}; \tag{10a}$$

$$\tilde{K}_\lambda(\mathbf{x}, t) = tr(\vec{\mathbf{E}}) * tr(\overleftarrow{\mathbf{E}}); \tag{10b}$$

$$\tilde{K}_\mu(\mathbf{x}, t) = \vec{\mathbf{E}} * \overleftarrow{\mathbf{E}}. \tag{10c}$$

All of the terms appearing in eq. (10) were correctly described in Colombi *et al.* (2012), including the concept of backwards and forwards fields. Following an analogous procedure, we transform the waveform kernels $\tilde{K}_\mu$, $\tilde{K}_\rho$ and $\tilde{K}_\lambda$ to traveltime kernels in a continuous form:

$$\Delta t = \int_\Omega \Big[\delta\rho(x) K_\rho(\mathbf{x}) + \delta\mu(x) K_\mu(\mathbf{x}) + \delta\lambda(x) K_\lambda(\mathbf{x})\Big] \mathrm{d}\mathbf{x}^3. \tag{11}$$

As we compute wavefields that solve the elastic system using AxiSEM (Nissen-Meyer *et al.* 2014), we follow the computational outline introduced by Colombi *et al.* (2012) to calculate eq. (11). In this framework, it appears particularly convenient for us to compute the λ kernels K_λ, as this involves only scalar products. K_λ directly expresses the sensitivity to v_p, the P-wave speed, such that we can consider the traveltime sensitivity in the following simplified manner:

$$\Delta t = \int_\Omega \delta v_p(x)\, K_{v_p}(\mathbf{x})\, \mathrm{d}\mathbf{x}^3. \tag{12}$$

Other relationships between ρ, μ and v_s can be found, for instance, in Tromp *et al.* (2005). A few examples of K_{v_p} can be found, for instance, in Nissen-Meyer *et al.* (2014). In equivalence to the boundary sensitivity, by choosing a set of $i = 1, \ldots, N$ base functions, φ_i

and a 'voxel', eq. (12) reduces to a simple scalar product between two vectors:

$$\Delta t = G_i \, \delta v_i, \tag{13}$$

where δv_i is now the *P*-wave velocity perturbation in the voxel i, and $G_i = K_i \, V_i$, is obtained by quadrature (e.g. Quarteroni *et al.* 2007) of the ith voxel volume V_i. If our database contains $j = 1, \ldots, M$ traveltime observations that we collect in a residual traveltime vector $\Delta\mathbf{t}$, the linearized inverse problem takes the following (matrix) form:

$$\Delta\mathbf{t} = \mathbf{G} \cdot \delta\mathbf{v}, \tag{14}$$

where $\mathbf{G}$ is an $M \times N$ matrix, and each row G_i corresponds to one discrete sensitivity kernel (eq. 13). Analogous considerations relating to the differences between continuous versus discrete kernels that were explained in Section 3.2 apply in the volumetric case. The rectangular system in eq. (14) can now be inverted using least-squares, with a procedure analogous to eq. (7), and with the appropriate choice of norm and roughness damping. The discrete equivalence of eq. (9) can easily be rewritten by combining the results from Section 3.1, to obtain the following linear system, 'the simultaneous joint inversion':

$$\Delta\mathbf{t} = \mathbf{G}_{v_p} \cdot \delta\mathbf{v} + \mathbf{G}_r \cdot \delta\mathbf{r}. \tag{15}$$

In practice, each row of $\mathbf{G}_{v_p+r}$ will have a number of columns that amounts to the number of pixels plus the number of voxels; that is, the global number of unknowns. The joint volume boundary inversion grid is constructed such that the voxels at the level of the CMB will exactly match the pixels of the CMB. A dimensional analysis of the quantities in eq. (9), according to the expression of the forwards and backwards fields, shows that K_{v_p} is measured in s/m^4, while K_r is measured in s/m^3, and thus the quantities in eq. (15) are correctly scaled and lead uniquely to a traveltime anomaly (e.g. Liu & Tromp 2008; Colombi *et al.* 2012). However, the different magnitude of the sensitivity kernels can drive the inversion preferentially in one direction. A number of studies have considered this problem and have suggested various solutions. Kennett *et al.* (1988) and Greenhalgh *et al.* (2006) used a subspace method that seeks the solution within each parameter space. Hobro *et al.* (2003) proposed a simpler method that by scaling each class of model parameter, eliminates the discrepancy between the different types of sensitivity. We chose this latter approach, because it is more intuitive and it matches with our formulation of the joint sensitivity matrix. The method requires the calculation of a normalization factor n_c for each class of parameter (in our case $c = 1, 2$) using the following relationship:

$$n_c = \sqrt{\sum_{i,j} [K_c]_{ij}^2}, \tag{16}$$

where $[K_c]_{ij}$ are the entries corresponding to one type of sensitivity. For the volumetric part, this will be the entries in the blue and yellow regions, while those in green are for the boundary entries (Fig. 7). The new sensitivity matrix is then obtained by dividing each sensitivity kernel by the appropriate n_c. This procedure regularizes the gradient of the misfit function, thereby stabilizing the inversion.

4 INVERSION OF SYNTHETIC DATA

Combining the residual waveform data computed via the spectral element simulations using SPECFEM3D GLOBE, as described in Section 2.3, and the inversion scheme discussed in Section 3, using the axial-symmetric solver, we can test the performance of the inversion algorithm. We also compare it to ray-based inversions, and check the applicability of our method despite the distortions introduced by the presence of 3-D mantle structures. The amount of data used by each inversion, and the number of model parameters for each grid, are shown in Tables 1 and 2, respectively. An interesting aspect of this approach is the possibility to obtain images for different frequency bands. The residuals are calculated from synthetics at 20- and 10-s dominant periods, and the sensitivity kernels are computed accordingly. The sizes of the inversion grids are those described in Fig. 6. The best damping factor α_1 is adjusted using the *L*-curve (Fig. 5c) calculated on the variance reduction introduced in Boschi & Dziewonski (1999). For the synthetic inversion, we do no make use of roughness damping, hence $\alpha_2 = 0$. The choice of α_1 does not have an important role, because the model variations are small and the noise is negligible; nevertheless $\alpha \neq 0$ is requested to avoid matrix singularities. We will often relate the lateral extension of the anomaly in terms of the wavelength γ. The images of the model shown in this section are all calculated for a 3°-large pixel inversion grid.

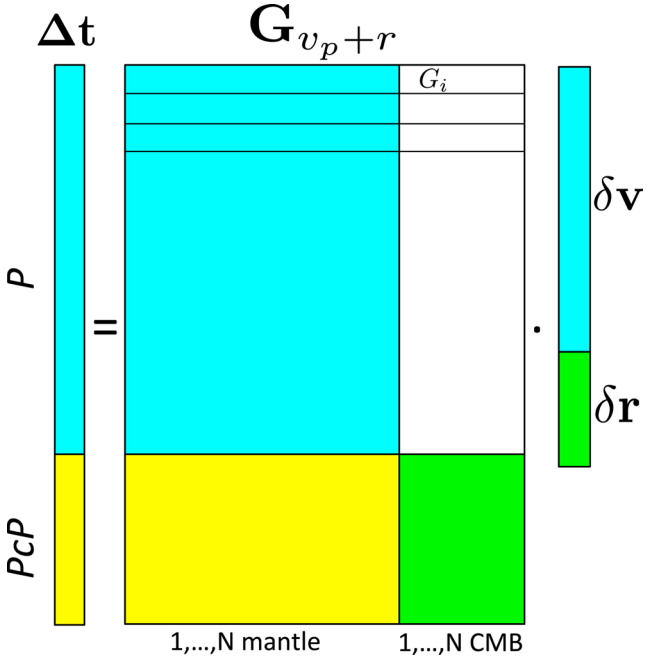
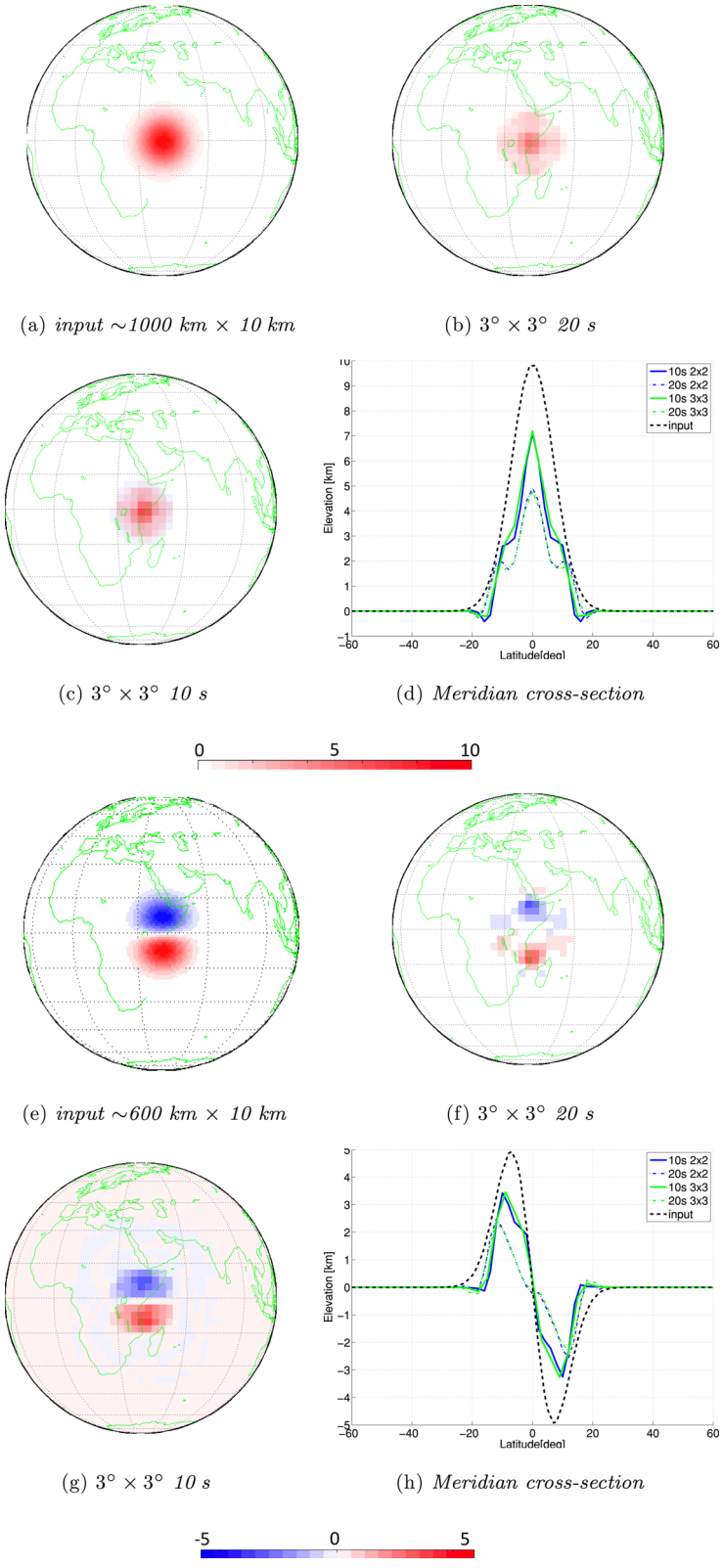


Figure 7. The joint sensitivity matrix. The block in blue represent the rows that are purely sensitive to the mantle, and are therefore associated to the *P* phase; note the 0 where the columns associated with the CMB are. The yellow represents the portion of the sensitivity associated with the mantle for *PcP*, and the green, those associated with the CMB.

4.1 How different phases sense the CMB

4.1.1 CMB topography with ScS/PcP

The inversion of *PcP* and *ScS* are performed in a similar fashion, and for reasons of brevity, we show only the results obtained for *ScS*, which are more easily recorded. Fig. 8 presents the inversion results for the two different anomalies, with medium length ($\sim 11\gamma$ for ~ 10 s; $\sim 4\gamma$ for 20 s) and small length ($\sim 6\gamma$ for 10 s; $\sim 3\gamma$ for 20 s) scales. A vertical cross-section going through the topography peak along a meridian (Figs 8d and h) helps to compare the inversion



(a) *input ∼1000 km × 10 km* (b) 3° × 3° *20 s*

(c) 3° × 3° *10 s* (d) *Meridian cross-section*

(e) *input ∼600 km × 10 km* (f) 3° × 3° *20 s*

(g) 3° × 3° *10 s* (h) *Meridian cross-section*

Figure 8. Synthetic inversion results using the *ScS* phase. (a) Input model inserted in SPECFEM3D GLOBE. (b) Recovered pattern using 3°-spaced pixels at a 20-s dominant period. (c) As for (b), but at a 10-s dominant period. (d) Value of the topography taken along the meridian cross-section (where the topography is centred) for the different dominant period/pixel size. (e–h) As for (a–d), but for a different input model. The colour bar is in kilometres.

output with the reference. The models with the larger lateral length scale are sufficiently reconstructed for the 20-s case, despite the expected amplitude reduction (Figs 8a–d); that is, the magnitude of the topography/depression is smaller. The second input model has a shorter length scale, and the reconstruction is acceptable when 10-s kernels are used (Fig. 8g). The models with 20-s kernels are poor in amplitude and lateral resolution; furthermore, the topography peaks are slightly misplaced.

4.1.2 CMB topography with PKP

The *PKP* branches have been used in several studies to constrain the CMB topography (Morelli & Dziewonski 1987; Soldati *et al.* 2003, 2012) or the structure of the lower mantle (Lay & Garnero 2004; Vanacore *et al.* 2010). These are characterized by double-sided sensitivity; that is, on the entry and the exit points. The plots of the sensitivity kernels show a first Fresnel zone of a size that is comparable with that of *PcP* (see Colombi *et al.* 2012, for sample images). Fig. 2(a) explains why we used only synthetic events located right on top of the topography. The bulk of the measurements is represented by *PKIKP*, because only few *PKPab*/*bc* residuals have been obtained (see explanations in Section 2.3).

Fig. 9 presents the results for the inversion using the second input model of the *ScS*/*PcP* case. We used a global map instead of a regional map to determine whether mapping on the exit point occurs. Positive and negative anomalies with a smaller lateral wavelength-to-length-scale ratio are more difficult to retrieve at 20 s, and artefacts are introduced, while good results are ensured at the 10-s dominant period, with an overall accuracy similar to *ScS*/*PcP* (Figs 9b–e). The coverage is good enough, such that almost nothing is mapped on the exit side of the *PKIKP* branch, except for a weak signal in Fig. 9(b). The settings of this test resemble regional studies like Vanacore *et al.* (2010), for instance, who used observations from single or only a few earthquakes, assuming good sensitivity of this phase over a relatively dense and closely spaced array of receivers.

4.1.3 CMB topography with P_{diff}

The core-diffracted P_{diff} waves do not carry as much information for topography as in the case of *ScS*/*PcP*. We nearly doubled the lateral length scale ($\sim 18\gamma$ for 10 s) of the anomaly to obtain usable imprints from the cross-correlation. In spite of this, the residuals can be measured only for the 10-s dominant period waveform. The kernels and inversions are therefore restricted in this case. The tomographic maps in Fig. 10 show poor reconstruction. This conclusion can be drawn *a priori* by inspecting the magnitude of the residual traveltimes,which is less than a tenth of the magnitude than the reflected or transmitted phases can provide (Colombi *et al.* 2012). The impossibility of reproducing a topography of a significant lateral length scale and consistent magnitude for higher dominant periods prevents us from making frequency-dependent comparisons. Our claim for this lack of sensitivity is:

(i) the non-negligible sensitivity surface is significantly larger than that of the reflected phases;
(ii) the wave front healing effects cancel the effects of the anomaly, as documented in Malcolm & Trampert (2010).

We conclude that the core-diffracted waves are not suited for CMB topographic inversions, because they do not provide sufficient sensitivity to the topography, in agreement with Colombi *et al.* (2012).

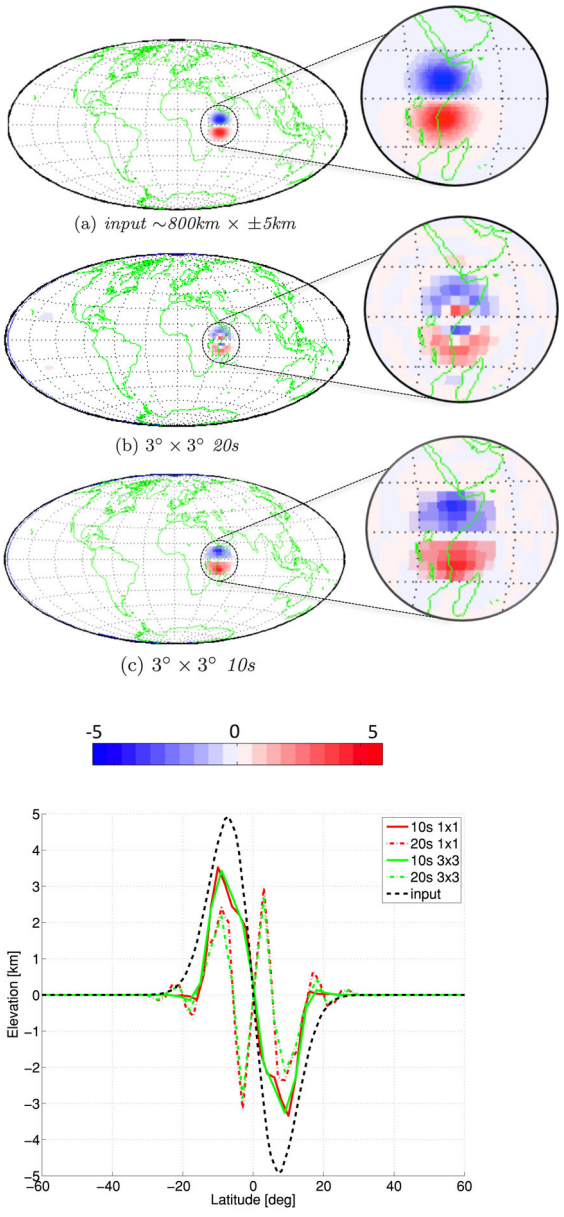
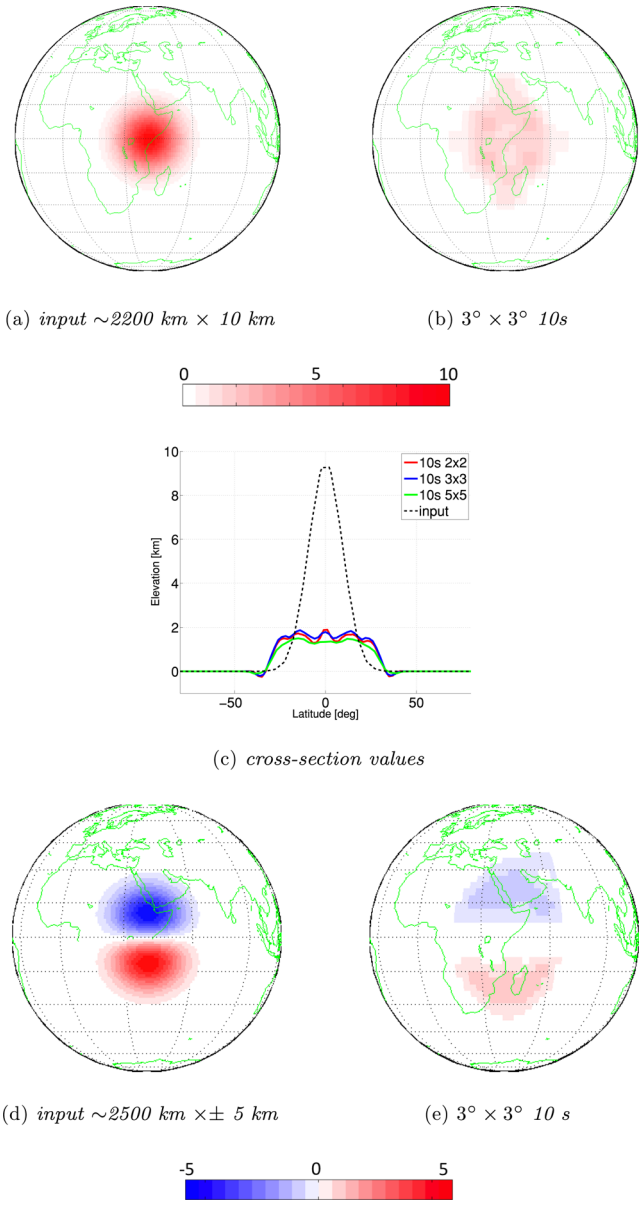


Figure 9. Topography inverted from the *PKP* branches. (b) Inversion output for 20-s dominant period and different grid parametrization using *PKIKP* recorded by the layout in Fig. 2 for the topography in (a). (c) As for (b), but for a dominant period of 10 s. (d) Cross-section value along the meridian passing through the centre of the topography, which shows how different parametrization/dominant periods perform. The colour map is in kilometres. To determine whether something was mapped on the exit point, we used global projection, rather than regional projection.

4.2 Finite frequency versus ray theory

Since cross-correlation delay times have been found to remain linear even in earth models characterized by relatively sharp heterogeneity (e.g Mercerat & Nolet 2013), it is worthwhile to compare our finite-frequency results to those of ray theory. We selected our shortest period data set (10 s) for one of the topography models available, and we use the residual traveltime for the ray theory code developed by Soldati *et al.* (2012), which determines the CMB topography models. Their inverse problem has been parametrized in a similar



(a) *input ∼2200 km × 10 km* (b) 3° × 3° *10s*

(c) *cross-section values*

(d) *input ∼2500 km ×± 5 km* (e) 3° × 3° *10 s*

Figure 10. Inversion results using the P_{diff} phase. (a) Input model inserted in SPECFEM3D GLOBE. (b) Recovered pattern using 3°-spaced pixels at a 10-s dominant period. (c) The value of the topography taken along the meridian cross-section (where the topography is centred) for the different dominant pixel sizes. (d and e) As for (a and b), but for a different input model.

fashion, but the iterative solver was LSQR of Paige & Saunders (1982). The results show that ray theory reconstructs the anomaly relatively well, provided that the inversion grid ensures sufficient coverage. Looking at the hit count table in Fig. 11(d), we see that coverage is missing in some areas, and the corresponding output map is biased. This does not happen with finite frequency, because the smoothing effect of the kernels produces smoother maps also where the ray coverage is not perfect. This can be deduced by looking at the L_1 column norm of $\mathbf{G}$ in Fig. 5(b), where the coverage with finite frequency appears more uniform (considering the same data set).

This suggests that finite frequency might perform better than ray theory on smaller pixel grids, or in other words, that finite frequency with the same parametrization requires less data to similarly constrain the model space.

4.3 Influence of source–receiver distribution and density

The source–receiver layout used for the previous test represents an ideal case, which is rarely found in reality. However, our approach allows the resolution and coverage capacities of the kernels to be quickly verified. To do so, we now arrange the source–receiver layout in various ways.

4.3.1 Acquisition grid coarsening

We eliminate records from the *PcP* database to reach 2°, 4°, 6° and 8° spaced receivers. The *PcP* data were used in this case because the coverage with this phase was good, and so were the inversion results. In the interest of conciseness, we do not show the results for each model, but rather measure the L_2 norm difference from the input model:

$$\chi = \sum_N \delta\mathbf{r}^2_{\text{input}} - \delta\mathbf{r}^2, \tag{17}$$

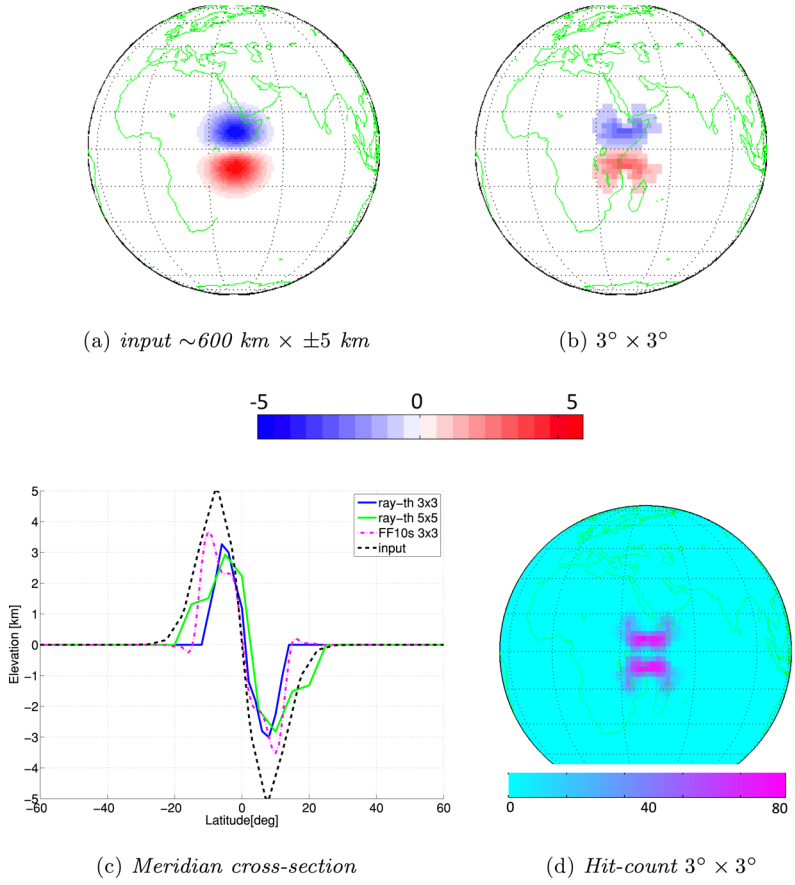
where the sum is intended over the inversion parameters $\delta\mathbf{r} = \delta r_i$ with $i = 1, \ldots, N$, and the input model is discretized over the same grid. We show the result only for the topography in Fig. 8(e), as this represents the most challenging case. Fig. 12 shows on the y-axis the value of χ normalized over the initial misfit χ_0, and, therefore, the differences are interpreted in a relative sense. The smoothing effect of the kernels helps to maintain a stable inversion; however, for more than 4°, the solution is not satisfactory any more, as the magnitude and shape, in particular, are lost. The number of observations is drastically reduced, and the stations with dominant Δt do not enter into the inversion. The degradation is slightly more pronounced when the grid used has a smaller pixel, as shown by Fig. 12(b).

4.3.2 USArray-like distribution

By virtue of symmetry, we can change our database; as it was covering North America, and we can select the stations close to their actual USArray positions. We do not have the same flexibility for the earthquake distribution, hence we select a subset, as shown in Fig. 13. The input anomaly is the one in Fig. 8(a), and it represents the most challenging test, using 10-s data. Although they are worse than in Fig. 8, the shape and magnitude of the anomalies are relatively well reconstructed. Note that this was possible only using ∼500 measurements, hence differences can be attributed to the reduced coverage of the data set and to a lack of cross-paths induced by the prevalent south–north orientation of the source–receiver pairs. In practice, a receiver distribution like USArray should be able to reconstruct such features even better, because a larger number of earthquakes covering a broader azimuth and area is available.

5 SEISMIC DATA INVERSION

We now use our algorithm to invert the cross-correlated traveltimes collected by Ritsema *et al.* (2011).This database includes ∼35 000 *SKSac*, ∼18 000 *PKIKP* and ∼9000 *ScS* phases. The number of these last observations is too small to be used for an inversion of *ScS* data only map, and therefore this is not shown. The data were relocated and corrected for ellipticity. The crustal corrections were



(a) *input* ∼*600 km* × ±5 *km* (b) 3° × 3°

(c) *Meridian cross-section* (d) *Hit-count* 3° × 3°

Figure 11. Ray theory inversion test. (a) Input model as used in Figs 8 and 9. (b) The ray theory model. The colour bar is in kilometres. (c) Meridian cross-section. (d) The hit count map associated with each inversion parametrization.

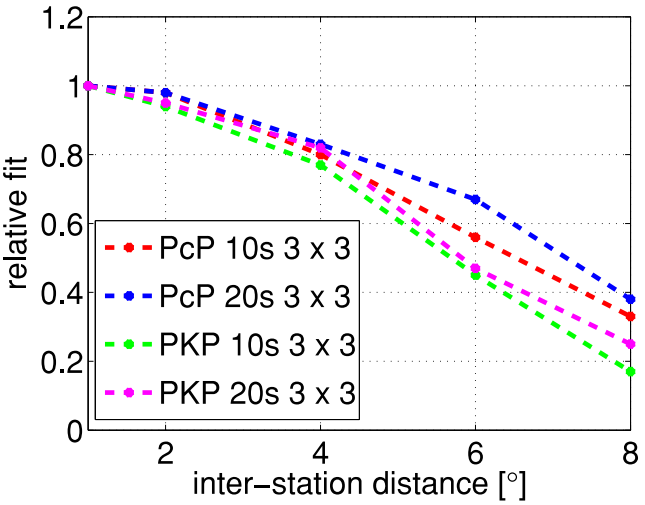
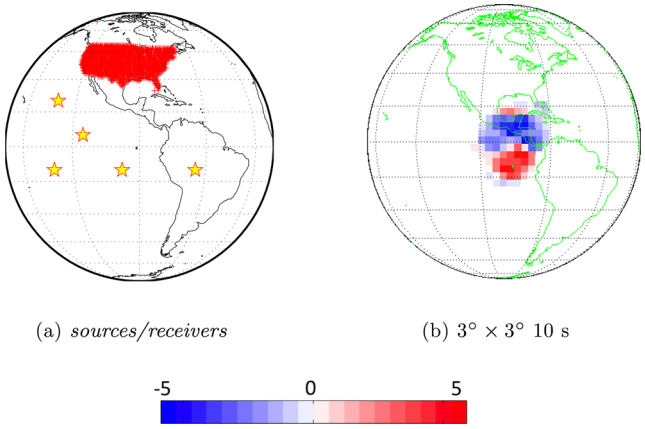


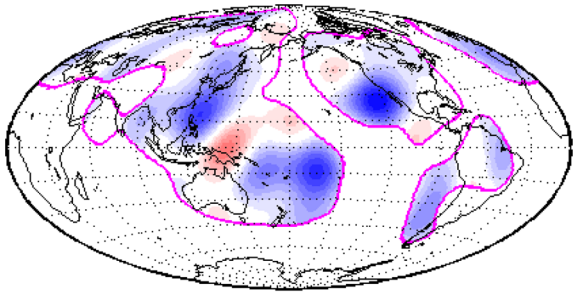
Figure 12. Plot that depicts how the input model is reconstructed when we turn off an increasing number of stations. Each colour corresponds to a phase and a dominant period. The curves are computed for a 3° × 3° pixel grid, for the input model of Fig. 8(e).



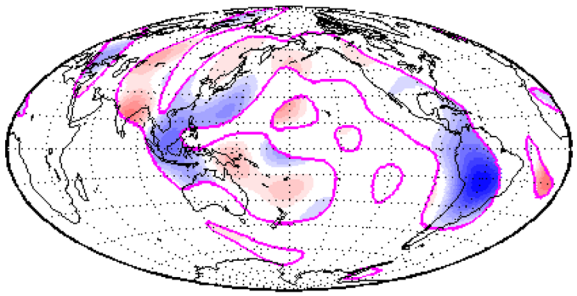
(a) *sources/receivers* (b) 3° × 3° 10 s

Figure 13. USArray-like acquisition grid. (a) Acquisition grid and the earthquake distribution used for this test. (b) Output model that is produced by inverting the data recorded from fictitious USArray stations. The input model is that of Fig. 8(e).

based on the model from Bassin *et al.* (2000), and were added to the residual traveltime as static corrections. The regularization scheme is based on norm and roughness damping to stabilize the solution and to smooth out the variations. We base the choice of α_1 and α_2 on the *L*-curve analysis introduced in Fig. 5. The CMB topography maps in Fig. 14 are combined with coverage maps that show the regions where there is sufficient data coverage, and thus the results are more reliable. The coverage is estimated using the L_1 norm of the columns of **G**. Thresholds are determined upon the results of the synthetic inversions. We filter the spherical harmonic spectrum of the tomographic maps to eliminate noise at $l > 10$, where l is the degree. Mantle corrections provided by the author and our corrections implemented using the body waves in the database of Ritsema *et al.* (2011) gave similar results. We used the former to correct for mantle structures, such that the residual traveltimes



(a) $SKSac$



(b) $PKIKP$

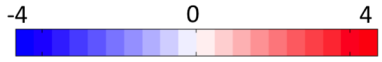


Figure 14. CMB topography maps from S40RTS cross-correlated traveltime data on a 3° × 3° grid filtered up to a spherical harmonic degree of 10. The contour lines isolate the regions where the data ensure sufficient coverage. (a) Topography map obtained using *SKSac* data, and its relative data coverage. The damping used is: $\alpha_1 = 3$ and $\alpha_2 = 3$. (b) Topography map obtained using *PKIKP* data, and its relative data coverage. The damping used is: $\alpha_1 = 1$ and $\alpha_2 = 4$. The colour bar for the topography maps is in kilometres.

should then be approximately sensitive only to the CMB. The effects of the outer core on the *PKIKP* and *SKS* are assumed to be negligible. The sensitivity kernels for the *S*-data are computed for a dipole moment tensor at 18 s located at varying depths according to 20-km bins. The same period, but an explosive source, is used for the *P* data. The vertical component of the velocity seismogram was used for the *SKS* and *PKIKP*. We first observe that the coverage is sufficient only along the Pacific rim. In the illuminated regions, we recognize some similarities with other models (e.g. Soldati *et al.* 2012). The presumed presence of deep subduction results into a depressed CMB (blue areas), especially in the *PKIKP* model. The difference between Figs 14(a) and (b) indicate that resolution is limited, which might be explained by reduced data quality or by the difficulties inherent to the presence of heterogeneity throughout the mantle, as discussed in the next section.

5.1 3-D mantle and topography

A classic way to account for the mantle effect on the traveltime is to eliminate its contribution to the total traveltime anomaly, as described by eq. (1) (static corrections). These corrections are calculated from an *a priori* 3-D mantle model that should represent the actual mantle as faithfully as possible (see also Section 2.4).

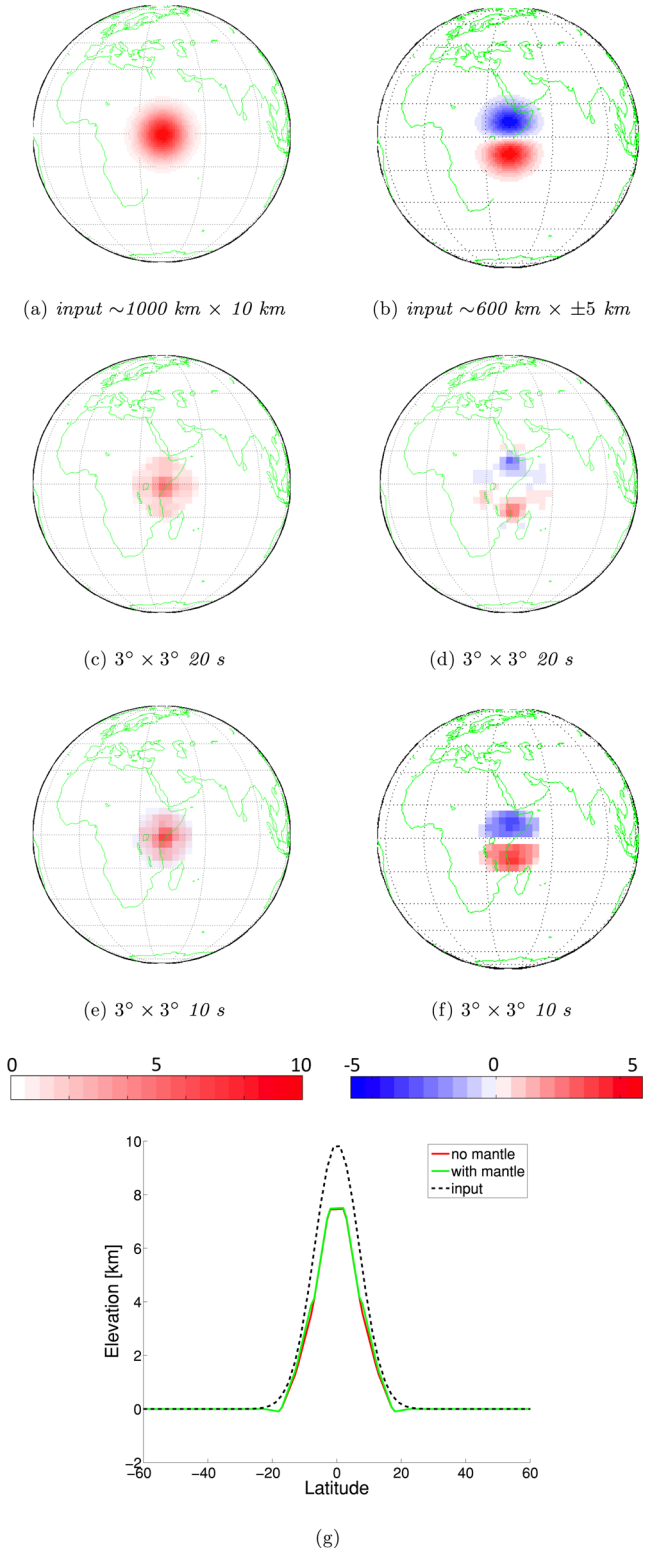
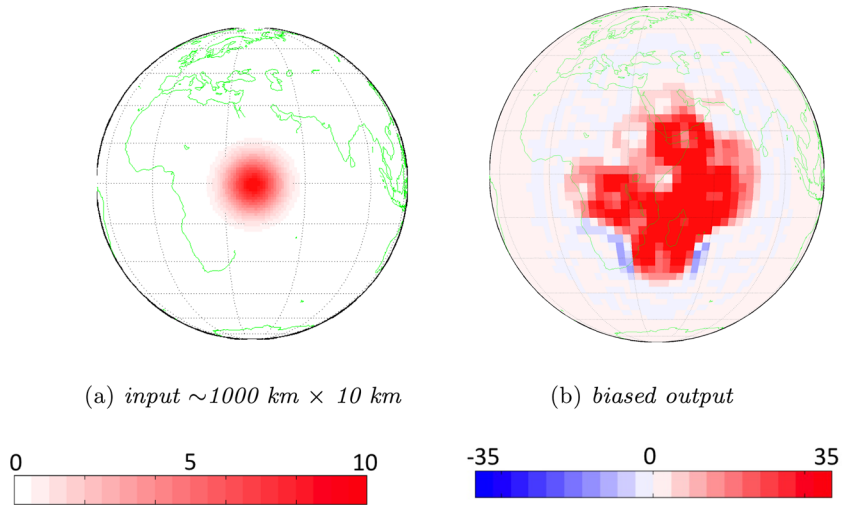


Figure 15. CMB inversion with the 3-D mantle correction according to eq. (1). (a and b) Input models. (c and d) Tomographic output for 20-s data and (e and f) the output model for 10-s data. (g) The meridian cross-section that compares the results with and without the mantle effects.

If this is the case, and if the finite frequency effects are taken into account for both the mantle and the CMB, the mantle effect can be ruled out, as as shown in Fig. 15. For this example, we used the synthetics corrected for the S40RTS mantle in Fig. 4. To



(a) *input ∼1000 km × 10 km* (b) *biased output*

Figure 16. (a) Input model. (b) Output model at 20 s after the incorrect mantle correction; that is, the mantle heterogeneity is underestimated by a factor of 2. The colour bars are in kilometres.

estimate the errors that inaccuracies in the mantle model would cause in our CMB maps, we conducted a test with the exaggeration of the mantle bias. The synthetics are those associated with the S40RTS model, but the mantle correction is now incorrectly calculated after dividing the S40RTS data amplitude by 2 for eq. (1). The result is shown in Fig. 16, where the input model is compared with the biased output model (up to a factor 4 bias). Artefacts appear because the mantle is mapped into the CMB. The error can be caused not only by a biased mantle model, but in a finite frequency context, also from the incorrect calculation of the frequency-dependent traveltime misfit by cross-correlation. In other words, the mantle correction applied to the CMB data must be calculated for the same frequency for eq. (1) to be valid. Probably, the only valid approach to tackle this problem of the mantle correction is through a mantle plus CMB joint inversion, which is further discussed in the next section.

6 OUTLOOK FOR JOINT VOLUMETRIC-BOUNDARY INVERSION

We next use the joint sensitivity that was introduced in Section 3.3, to investigate how simultaneous joint inversion can improve the CMB maps obtained previously. A qualitative way to determine whether the delay due to the mantle on the CMB sensitive phases can be isolated, is to reverse the modelling residual traveltimes due to both the mantle anomalies and the CMB topography. Focusing only on the *P* and *PcP* phases, we computed the sensitivity kernels for the source–receiver pairs that experienced similar sensitivity through the mantle in a simplified scenario. We computed $\mathbf{G}_{v_p}$ for the source–receiver layout of Fig. 1(d). We selected the source–receiver pairs for which the epicentral distance was between 65° and 80°. For this distance range, the volumetric sensitivity of the *P* and *PcP* phases is similar for the first ∼1000 km depth (e.g. Liu & Tromp 2008), and therefore an anomaly located in this area should be sensed in a similar way. This idealized scenario is represented in Fig. 17. A rapid anomaly that is inserted between 100 and 900 km in depth and 45° × 45° in size is located directly above the positive CMB topography of the type in Fig. 1(e), at a 30-km depth and ∼6000-km large. The receiver grid and the fast anomaly are centred

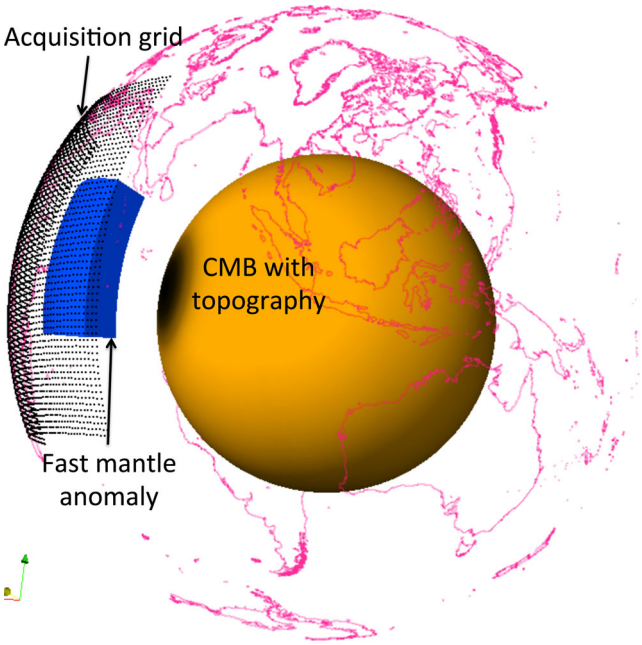


Figure 17. Portrait of the rapid mantle anomaly ($\delta v_p/v_p = 0.02$) in the mantle (blue section) located above the positive topography anomaly (30 km). The black dots are the receivers that are distributed according to the layout of Fig. 1(d). The events (not shown) also follow Fig. 1(d).

above the topography peak. In a second example, we move the fast anomaly into the deep mantle between 1900 and 2700 km.

Using the procedure in Sections 3.3 and 4, we compute K_{v_p} for the *P* phases, and $[K_{v_p}, K_d]$ for the *PcP* phases, which we then project onto the inversion grid. The normalization factor that appears in eq. (16) is not necessary, because we only study the forward problem. The computations were done at 40 s, to reduce the computational effort in calculating the volumetric kernels. To compensate for the reduction of sensitivity that longer period waves have, we increase the size and extension of the topography anomaly according to the results obtained by Colombi *et al.* (2012) during synthetic and reverse modelling tests. The distance range was sufficiently large for an accurate projection onto a 3° × 3° inversion grid of

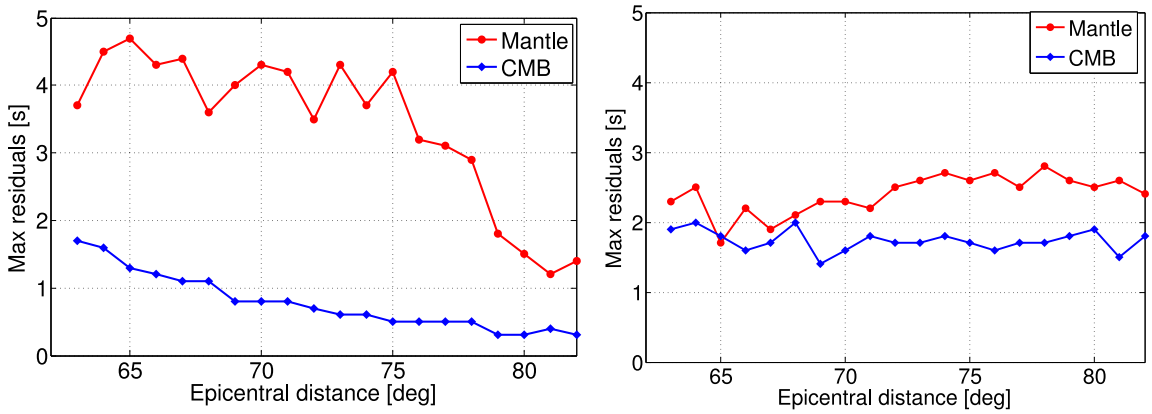


Figure 18. (a) Summary traveltime residuals for the mantle only and the topography calculated for equal source–receivers only for the *P* (red) and *PcP* (blue) kernels, as these latter would feel only the CMB. The anomaly was located in the upper mantle. (b) As for (a), but for a wave speed anomaly located in the lower mantle.

the *P* kernels. The source–receiver layout chosen ensures sufficient coverage along all of the azimuths surrounding the anomaly for a total of ∼4500 observations of the *P* and *PcP* phases. The choice of using the differential *P*–*PcP* sensitivities to correct for the overlying structures is known in seismic tomography. We claim that although this method is good for crust and upper-mantle heterogeneities, because the sensitivity is similar (both using ray theory and finite frequency), it might fail for anomalies located in the lower mantle, where the sensitivity patterns differ significantly. We next model the traveltime residuals Δt using eqs (13) and (3), for each source–receiver pair. The net effect of the mantle Δt_P captured by the *P* phase is subtracted from the Δt_{PcP}, which contains both. The result is a proxy for the residual due to the CMB. As there are a large number of source–receiver pairs that sample anomalies in various ways, it is important to filter the results using the summary data, like the maximum or minimum value over the entire data set. Both the rapid anomaly and the positive topography produce negative residuals, anticipating the arrival time of the phase. This helps to avoid ambiguity in our analysis. Over all of the possible source–receivers pairs (that are within the distance range), we calculate the maximum value of $\Delta t_{PcP} - \Delta t_P$ and Δt_P for the same distance ranges. The distance range and the kernels are calculated using 1° bins. In Fig. 18(a), we plot the values of Δt_P (red) and $\Delta t_{PcP} - \Delta t_P$ (blue) as functions of the epicentral distance. This latter represents the traveltime residual due to the topography. It shows a monotone decrease, which agrees with the average traveltime anomaly due to the CMB topography only for each epicentral distance bin. The traveltime residuals generated from the mantle (Fig. 18a, red) feature stronger oscillations, due to the irregular sampling of the velocity anomaly. These oscillations do not appear to propagate in the blue line of Fig. 18(a). On the other hand, Fig. 18(b) shows that when the velocity anomaly is located in the lower mantle, this method fails, because the sensitivity in that region differs too much. Bearing in mind that the K_{v_p} component of the *P* and *PcP* sensitivity kernels are very similar only in the upper mantle, we can conclude that the monotone trend of the blue line in Fig. 18(b) (the CMB topography effect) is a good indicator that the mantle contribution can be ruled out from the combined effects of the mantle heterogeneities and the CMB topography. Therefore, a joint inversion featuring the data with good coverage of the whole mantle can successfully capture the volumetric effect and prevent its propagation in the boundary term. The validity of such a test needs, however, to be verified, together with the coverage required in the presence of a complex mantle and CMB structures with various length scales.

7 CONCLUSIONS

Using a comprehensive synthetic database, we have tested the theory and methodology for constraining the boundary topography developed in Colombi *et al.* (2012), and we have applied this to actual global seismic data sets of both compressional and shear waves. With respect to Colombi *et al.* (2012), our synthetic database is more complete, in that it contains waveforms for *PKP* and *ScS* arrivals associated with a set of different CMB and mantle models. To account for complex laterally varying mantle models, part of the synthetics database is obtained using the 3-D background S40RTS model by Ritsema *et al.* (2011), down to a period of 10 s. Using this latter setting, we have confirmed that static corrections derived from 3-D models are valid, and provided that the mantle model is close to the actual mantle, the effects of boundary topography are separable from those of the mantle. Exploring the synthetic data, we found that the magnitude of the traveltime anomaly depends on the frequency and it is as small as a tenth of the traveltime anomaly caused by mantle heterogeneities if any of the current CMB topography models are used. This leads us to conclude that CMB topography residuals are likely to be strongly biased by overlying structures if one does not accurately account for the two contributions explicitly.

In the analytical Section 3, we outlined the formulation of the inverse problem in a least-squares sense, using roughness and norm damping and starting from the results in Colombi *et al.* (2012). We completed this section by providing the expressions for the mantle sensitivity kernels, and we showed how these can be merged with boundary kernels, to set up a simultaneous joint inversion for compressional wave speed and CMB topography. A normalization factor between the two classes of parameters is, however, necessary.

The results of the synthetic inversions for the *PcP*, *ScS* and *PKP* show that by using the correct frequency contents, the pattern of the input model is retrieved successfully. The magnitude is also well reconstructed, although the shortest features results smoothed out. The input model for P_{diff} cannot be retrieved in a satisfactory manner. We highlighted the importance of choosing the correct parametrization, in agreement with the wavelength of the anomaly we anticipate, and consequently the dominant period at which the kernels are calculated. These parameters have to accord to the data coverage, which can be verified using the L_1 norm of the columns of $\mathbf{G}$, as shown in Sections 3 and 5. According to our results, at a fixed frequency range, the output from *PcP*, *ScS* and *PKP* should be almost the same. Ray theory does not perform much worse than finite frequency for our specific setting, although some issues with the

model coverage emerged when inverting one synthetic data set. Using an excessively fine parametrization is discouraged, as this gives rise to spurious small-scale features, as shown in Section 4.2. Varying the source–receiver layout to approximate the actual coverage shows that rather than the position of the earthquakes, a dense acquisition grid similar to that installed by USArray is more important. With the dominant period used in this study (10–20 s), the structures that appear in current models should be reconstructible.

Starting from a 0-topography CMB model, we separately inverted cross-correlated traveltime data for *PKIKP* and *SKSac* from Ritsema *et al.* (2011), to obtain maps that are characterized by non-uniform coverage. We find that areas of high/low topography are to some extent correlated with areas of low/high velocity in the lowermost mantle, confirming the idea that CMB topography might be controlled by mantle flow (Boschi *et al.* 2013). Yet, important discrepancies between the two maps of Fig. 14 indicate a lack of data resolution or a biasing effect due to mantle correction. With a synthetic inversion, we show how this can limit the capacity to resolve the CMB topography.

With the aim of setting up a simultaneous joint inversion in a future study, we confirmed the capacity of the mantle sensitivity kernel for capturing the traveltime anomaly due purely to mantle heterogeneities, with respect to the joint effects of topography and the mantle. The results from the reverse modelling show that in a simple scenario with just a layer of negative anomaly in the upper mantle and a simple CMB topography, the delay times due to the mantle calculated from the *P* kernels are correctly extracted from those of the *PcP* kernels that contain both. Conversely, if the anomaly is located in the lower mantle, the topography signature is less accurately retrieved, due to the sensitivity trade-off with deep-mantle heterogeneities. The results from this test suggest that a joint inversion that features mantle and CMB sensitive data and their associated sensitivity kernels will improve the visibility of the CMB topography. A sensitivity analysis using the joint method will be the basis of a future study. In particular, it will aim to set a threshold on the minimum topography resolvable using various synthetic data sets. In parallel to that, we need to collect an adequate data set of cross-correlated traveltimes that ensure both good CMB coverage, and good sampling of the upper and lower mantle regions.

ACKNOWLEDGEMENTS

We thanks Jeroen Ritsema for giving us access to S40RTS database, Guust Nolet for the numerous discussions and advises, Steve Della Mora for the calculation of ray theory residuals and Ludwig Auer for helping with mantle corrections. Computations were undertaken on the Brutus Cluster at ETH Zurich and at the Swiss National Supercomputing Center (CSCS). We acknowledge funding from QUEST, an international training network by the European Commission.

REFERENCES

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.

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., Weemstra, C., Verbeke, J., Ekström, G., Zunino, A. & Giardini, D., 2013. On measuring surface-wave phase velocity from station-station cross-correlation of ambient signal, *Geophys. J. Int.,* **192,** 346–358.

Colombi, A., Nissen-Meyer, T., Boschi, L. & Giardini, D., 2012. Seismic waveform sensitivity to global boundary topography, *Geophys. J. Int.,* **191,** doi:10.1111/j.1365-246X.2012.05660.x.

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

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

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

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

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.,* **175**(2), 665–685.

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.

Forte, 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.

Garnero, E.J., 2006. Interrogating the Deep Earth with USArray, *IRIS Newsletter,* **3,** 1013–1016.

Golub, G.H. & Van Loan, C.F., 1996. *Matrix Computations,* 3rd edn, Johns Hopkins University Press.

Greenhalgh, S.A., Bing, Z. & Green, A., 2006. Solutions, algorithms and inter-relations for local minimization search geophysical inversion, *J. Geophys. Eng.,* **3**(2), doi:10.1088/1742-2132/3/2/001.

Hobro, J.W.D., Singh, S.C. & Minshull, T.A., 2003. Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltime data, *Geophys. J. Int.,* **152**(1), 79–93.

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.,* **1**(145), 77–96.

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,* Vol. 12, pp. 97–107, eds LeMoul, J.L., Smylie, D.E. & Herring, T., AGU Geophysical Monograph, IUGG.

Kennett, B.L.N., Sambridge, M.S. & Williamson, P.R., 1988. Subspace methods for large inverse problems with multiple parameter classes, *Geophys. J.,* **94**(2), 237–247.

Koelemeijer, P.J., Deuss, A. & Trampert, J., 2012. Normal mode sensitivity to earths D layer and topography on the core-mantle boundary: what we can and cannot see, *Geophys. J. Int.,* **190,** 553–568.

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.

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.

Lay, T. & Garnero, E.J., 2004. Core-mantle boundary structures and processes, in *The State of the Planet: Frontiers and Challenges in Geophysics,* Geophysical Monograph 150, Vol. 19, eds Sparks, R.S.J. & Hawkesworth, C.J., IUGG.

Li, X.-D., Giardini, D. & Woodhouses, J.H., 1991. Large-scale three-dimensional even-degree structure of the earth from splitting of long-period normal modes, *J. geophys. Res.: Solid Earth,* **96**(B1), 551–577.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.,* **174,** 265–286.

Malcolm, A.E. & Trampert, J., 2010. Tomographic errors from wave front healing: more than just a fast bias, *Geophys. J. Int.,* **185,** 385–402.

Mercerat, E.D. & Nolet, G., 2013. On the linearity of cross-correlation delay times in finite-frequency tomography, *Geophys. J. Int.,* **192**(2), 681–687.

Morelli, A. & Dziewonski, A.M., 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. & Dahlen, F.A., 2008. A 2-D spectral-element method for computing spherical-earth seismograms–II. Background models, *Geophys. J. Int.,* **174,** 873–888.

Nissen-Meyer, T., van Driel, M., Stähler, S., Hosseini, K., Hempel, S., Auer, L., Colombi, A. & Fournier, A., 2014. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth Discuss.,* **6,** 265–319.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *TOMS,* **1**(8), 43–71.

Panning, M., Capdevill, Y. & Romanowicz, B., 2009. Do first order 3D born finite-frequency kernels improve modeling of surface waveforms? *Geophys. J. Int.,* **177,** 161–178.

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.

Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.,* **186**(2), 721–739.

Quarteroni, A., Sacco, R. & Saleri, F., 2007. *Numerical Mathematics,* 2nd edn, Springer.

Richards, P.G., 1973. Calculation of body waves, for caustics and tunnelling in core phases, *Geophys. J. R. Astronom. Soc.,* **35**(1–3), 243–264.

Ritsema, J. & Van Heijst, H.J., 2002. Constraints on the correlation of p- and s-wave velocity heterogeneity in the mantle from p, pp, ppp and pkpab traveltimes, *Geophys. J. Int.,* **149**(2), 482–489.

Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltimes, and normal-mode splitting function measurements, *Geophys. J. Int.,* **184**(3), 1223–1236.

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.

Soldati, G., Boschi, L. & Forte, A., 2012. Tomography of core-mantle boundary and lowermost mantle coupled by geodynamics, *Geophys. J. Int.,* **189,** 730–746.

Soldati, G., Koelemeijer, P., Boschi, L. & Deuss, A., 2013. Constraints on core-mantle boundary topography from normal mode splitting, *Geochem. Geophys. Geosyst.,* **14**(5), 1333–1342.

Sze, E. & Van der Hilst, R.D., 2003. Core mantle boundary topography from short period PcP, PKP, and PKKP data, *Phys. Earth planet. Inter.,* **135,** 27–46.

Tanaka, S., 2010. Constraints on the coremantle boundary topography from P4KP-PcP differential travel times, *J. geophys. Res.,* **115,** B04310, doi:10.1029/2009JB006563.

Tape, C., Liu, Q. & Tromp, J., 2006. Finite-frequency tomography using adjoint methods, methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168,** 1105–1129.

Thorne, M.S., Lay, T., Garnero, E.J., Jahnke, G. & Igel, H., 2007. Seismic imaging of the laterally varying D″ region beneath the Cocos Plate, *Geophys. J. Int.,* **170,** 635–648.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Valenzuela, R.W. & Wysession, M.E., 1998. Illuminating the base of the mantle with diffracted waves, *Geodynamics,* **28,** 57–71.

Vanacore, E., Niu, F. & Ma, Y., 2010. Large angle reflection from a dipping structure recorded as a pkikp precursor: evidence for a low velocity zone at the core-mantle boundary beneath the gulf of mexico, *Earth planet. Sci. Lett.,* **293**(1–2), 54–62.

Vasco, D.W., Pulliam, R.J. & Johnson, L.R., 1995. Lateral variations in mantle velocity structure and discontinuities determined from *p*, *pp*, *s*, SS, and SS-SdS traveltime residuals, *J. geophys. Res.,* **100,** 24 037–24 059.

Vasco, D.W., Johnson, L.R. & Marques, O., 1999. Global Earth structure: inference and assessment, *Geophys. J. Int.,* **137**(2), 381–407.

Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC63–WCC76.

APPENDIX: TECHNICAL IMPLEMENTATION OF THE INVERSION WORKFLOW

The computation of the forward and backward quantities with the software of Nissen-Meyer *et al.* (2007b) and a wise optimization of the kernel calculation have made the inversion affordable. The waveform treatment to extract traveltimes, forward and backward field computations, kernel computations and matrix inversions were carried out on the 'Brutus' cluster at ETH Zurich. The crucial task was to optimize the kernel computations for thousands of source–receiver pairs, not through message passing (MPI), but leaving the parallelization to the batch system of the cluster. Each kernel computation was seen as a single job over the whole database; that is, exploiting another dimension of parallelism of the seismic inverse problem. This trick increases the throughput (2- to 3-fold) on that machine. This can be done as long as the memory required by the dynamic allocation does not exceed the RAM available on each CPU. Specifications relating to the kernels algorithm can be found in Colombi *et al.* (2012). Fig. A1 shows the conceptual scheme we put into place to carry out the inversions.

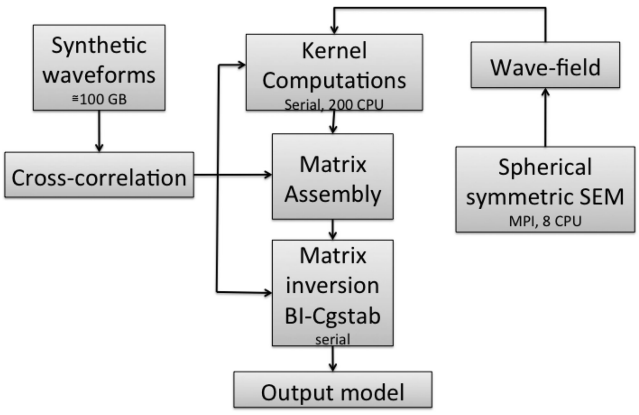


Figure A1. The inversion workflow set-up. The synthetic database is compressed to occupy a few hundreds of GByte. The data are processed on the cluster nodes, and the residuals for the desired source–receiver pairs are stored. The wavefield values are computed only once, and they are stored 'permanently' on the disk, and reused for every kernel, because the event characteristic does not change (see Colombi *et al.* 2012, for details). The kernel computation, is paralleled by the cluster scheduling system without the need for message passing.