Chapter 4
Interpreting Radial Anisotropy in Global and Regional Tomographic Models

Thomas Bodin, Yann Capdeville, Barbara Romanowicz and Jean-Paul Montagner

Abstract We review the present status of global and regional mantle tomography and discuss how resolution has improved in the last decade with the advent of full waveform tomography and exact numerical methods for wave-field calculation. A remaining problem with full waveform tomography is computational cost. This leads seismologists to only interpret the long periods in seismic waveforms and hence only constrain long-wavelength structure. In this way, tomographic images do not represent the true Earth, but rather a smooth effective, apparent, or equivalent model that provides a similar long-wavelength data fit. In this paper, we focus on the problem of apparent radial anisotropy due to unmapped small-scale radial heterogeneities (e.g., layering). Here, we propose a fully probabilistic approach to sample the ensemble of layered models equivalent to a given smooth tomographic profile. We objectively quantify the trade-off between isotropic heterogeneity and strength of anisotropy. The non-uniqueness of the problem can be addressed by adding high-frequency data such as receiver functions, sensitive to first-order discontinuities. We show that this method enables us to distinguish between intrinsic and artificial anisotropy in 1D models extracted from tomographic results.

Keywords Seismic tomography · Inverse theory · Bayesian inference · Monte Carlo methods · Seismic anisotropy

T. Bodin (✉) · B. Romanowicz
Berkeley Seimological Laboratory, UC Berkeley,
215 McCone Hall, Berkeley, CA 94720-4760, USA
e-mail: thomas.bodin@berkeley.edu

T. Bodin
Laboratoire de Géologie de Lyon: Terre, Planètes et Environnement, CNRS,
Université de Lyon 1, Ecole Normale Supérieure de Lyon, Villeurbanne, France

Y. Capdeville
Laboratoire de Planétologie et Géodynamique de Nantes, CNRS,
Université de Nantes, Nantes, France

B. Romanowicz · J.-P. Montagner
Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05 Paris, France

A. Khan and F. Deschamps (eds.), *The Earth's Heterogeneous Mantle*,
Springer Geophysics, DOI 10.1007/978-3-319-15627-9_4

4.1 Introduction

For more than thirty years, seismologists have imaged the earth's interior using seismic waves generated by earthquakes and traveling through different structures of the planet. A remaining challenge in seismology is to interpret the recovered Earth models in terms of physical properties (e.g., temperature, density, mineral composition) that are needed for understanding mantle dynamics and plate tectonics. For example, a region of slow wave speed can be either interpreted as anomalously warm, or rich in water, or iron.

Although seismic waves are sensitive to a large number of viscoelastic parameters as well as density, the mantle models constructed from seismic tomography are only parameterized with a few physical parameters, for example, average isotropic shear-wave velocity and radial anisotropy (e.g., French et al. 2013). This is because given the available information observed at the surface, there is not enough resolution to entirely describe the local elastic tensor. In addition to the limited number of resolvable elastic (and anelastic) parameters, there is also the question of spatial resolution, namely the smallest spatial scale at which heterogeneities can be imaged.

The number of independent elastic parameters that can be constrained is intrinsically associated with the level of spatial resolution. For example, it is well known that a stack of horizontal isotropic layers will be equivalent, at large scales, to a homogeneous anisotropic medium (Backus 1962). As we increase the scale at which we "see" the medium (the minimum period in the observed waveforms), we lose the ability to distinguish different layers, as well as the ability to distinguish between isotropy and anisotropy. The anisotropy observed at large scales may be artificial and simply the effect of unmapped fine layering. In other words, whether a material is heterogeneous (and described by a number of spatial parameters) or anisotropic (described by different elastic parameters) is a matter of the scale at which we analyze its properties (Maupin and Park 2014).

Therefore, there is a trade-off between spatial roughness and anisotropy when inverting long-period seismic data. By introducing anisotropy as a free parameter in an inversion, tomographers are able to fit seismic data with smoother models and fewer spatial parameters (Montagner and Jobert 1988; Trampert and Woodhouse 2003).

In this manuscript, we will first describe the issues that limit resolution in seismic imaging at regional and global scales (uneven data sampling, limited frequency band, data noise, etc), with a focus on the significance of observed seismic anisotropy and on the problem of distinguishing its different possible causes. Following Wang et al. (2013) and Fichtner et al. (2013a), here, we make the distinction between intrinsic anisotropy and extrinsic (i.e., artificial) anisotropy induced by structure. In the last section, we propose a method to separate these two effects in a simplified 1D case with vertical transverse isotropy (i.e., ignoring azimuthal anisotropy).

4.2 The Resolving Power of Regional and Global Seismic Tomography

It can be proven that if one had an unlimited number of sources, receivers, and an unlimited frequency band, one would be able to entirely describe an elastic medium from the displacement of elastic waves propagating through it and observed at its surface. That is, the function linking an elastic medium subjected to excitation by a source to the displacement measured at its boundaries is bijective. For detailed mathematical proofs, see Nachman (1988), Nakamura and Uhlmann (1994), and Bonnet and Constantinescu (2005).

However, in seismology, there are a number of elements that limit the resolving power of seismic observations, i.e., the ability to image structure. Firstly, the seismic records are limited both in time and frequency, and the number of sources and receivers is limited. Furthermore, there are a number of observational and theoretical errors that propagate into the recovered images. Finally, the earth is not entirely elastic, and seismic energy is dissipated along the path.

In this section, we give a brief description of these limiting factors which directly condition the level of resolution. Note that here, the phrase "level of resolution" or "resolving power" will be used in a broad sense and defined as the quantity of information that can be extracted from the data (the maximum number of independent elastic parameters or the minimum distance across which heterogeneities can be mapped). Here, we do not consider the resolution as it is mathematically defined in linear inverse theory and represented by a resolution matrix (e.g., Backus and Gilbert 1968; Aki et al. 1977), which, for example, does not depend on data noise or theoretical errors.

4.2.1 Different Seismic Observables

There are a multitude of ways of extracting interpretable information from seismograms. Due to practical, theoretical, and computational considerations, imaging techniques often only involve a small part of the seismic record. Different parts of the signal can be used, such as direct, reflected, and converted body waves, surface waves, or ambient noise. Different components of the signal can be exploited such as travel times, amplitudes, shear-wave splitting measurements, waveform spectra, full waveforms, or the entire wave-field (for comprehensive reviews, see Rawlinson and Sambridge 2003; Romanowicz 2003; Liu and Gu 2012).

Each observable has its own resolution capabilities. For example, analysis of converted body waves, now widely called the "receiver function", is used as a tool to identify horizontal discontinuities in seismic velocities (small-scale radial heterogeneities), but fails at determining long-wavelength anomalies. Conversely, surface wave measurements are sensitive to 3D absolute S wave velocities, but cannot constrain sharp gradients, and are poor at locating interfaces. Surface

wave-based imaging usually involves only the relatively low-frequency component of seismograms and is particularly effective in mapping the large-scale pattern of upper mantle structure. We will show how the seismic discontinuities that can be constrained with converted and reflected body waves are sometimes seen by surface waves as anisotropic structure.

Hence, the gaps between existing models can be described in terms of seismic wavelengths. The difficulty of assembling different databases with different sensitivities that sample the earth at different scales and the differences in the theory relating earth structure to seismic data of different nature have resulted in most models being based only on a limited portion of potentially available observations.

4.2.2 An Uneven Sampling of the Earth

One of the most important causes of poor resolution in seismic tomography is limited sampling of the volume of interest. In global seismic mantle tomography, there is no control on the distribution of the earthquake sources, which mostly occur at plate boundaries. Moreover, most receivers are located on continents, which cover only about one-third of the surface of the planet. This results in an uneven distribution of sources and receivers, especially in the southern hemisphere.

Traditional tomography relies primarily on the information contained in the travel times of seismic phases that are well separated on the seismic record: first arriving P and S body waves on the one hand and fundamental mode surface waves on the other. For the latter, which are dispersive, the measured quantity is the phase or group velocity as a function of period, in a period range accessible for teleseismic observations, typically ∼30 to ∼250 s.

The theoretical framework is typically that of infinite frequency ray theory for body waves, or its equivalent for surface waves, the "path average approximation" (PAVA) [see reviews by Romanowicz (2002), Romanowicz et al. (2008)]. Below, we briefly discuss how body and surface waves sample the earth differently, and then discuss how waveform tomography allows us to compensate for the non-uniform distribution of sources and receivers by exploiting more fully the information contained in each seismogram.

4.2.2.1 Body Wave Tomography

Because of the lack of stations in the middle of the oceans, body wave tomography based on first arrival travel times achieves best resolution in regions where the density of both sources and stations is high, typically in subduction zone regions around the Pacific Ocean and in the Mediterranean region (e.g., Bijwaard et al. 1998; Kárason and Van Der Hilst 2000; Fukao et al. 2001; Rawlinson et al. 2015). Much progress has been made in the last few years, owing to improvements in both quality and quantity of

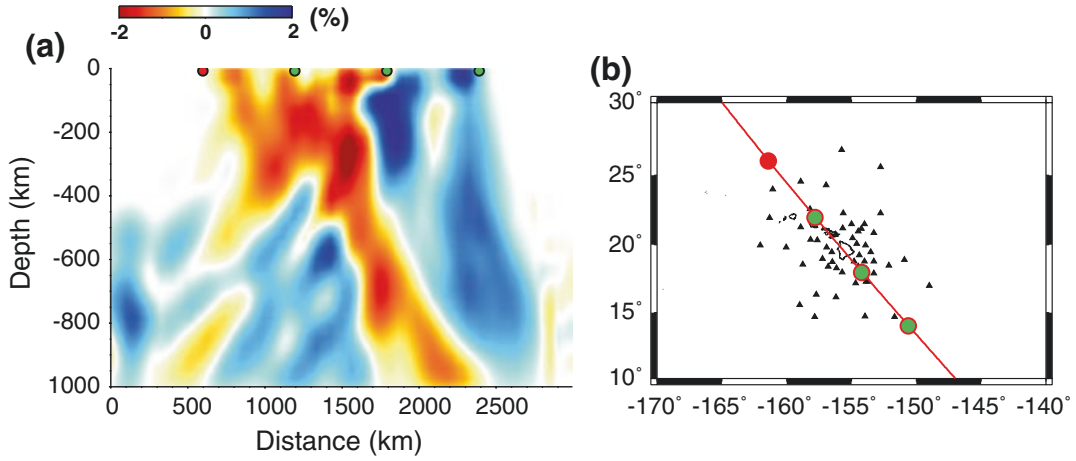

Fig. 4.1 Example of teleseismic body wave tomography under Hawaii with poor vertical resolution, i.e., vertical smearing. **a** Vertical cross sections (parallel to the Pacific plate motion) through the HW13 model (Cheng et al. 2015). **b** Locations and orientation of the cross section, along with the distribution of stations. This is a typical example of limited resolution due to poor data sampling. Modified from Cheng et al. (2015)

seismic data. Some technical improvements have also been made, such as the introduction of finite frequency kernels that take into account the sensitivity of the body wave to a broader region around the infinitesimal ray path (e.g., Dahlen et al. 2000). These improvements have led to increasingly high-resolution images in the last ten years indicating different behaviors of slabs in the transition zone, with some ponding on the 660 km discontinuity, and/or around 1000 km depth, while others appear to penetrate deep into the lower mantle (e.g., Li et al. 2008; Fukao and Obayashi 2013).

In other parts of the world, where only teleseismic data can be used, resolution in body wave travel time tomography depends strongly on the density of stations. In the oceans and in poorly covered continental regions, there is very poor vertical and horizontal resolution in the upper mantle, even when considering finite frequency effects, because of smearing effects due to the lack of crossing paths. In Fig. 4.1, we show an example of regional body wave tomography under Hawaii, where only teleseismic events originating at subducting zones around the Pacific are used (Cheng et al. 2015). Seismic rays arrive almost vertically under the array of stations, which results in poor vertical resolution as velocity anomalies are "smeared" along seismic rays. In this context, interpretation of the vertical plume-like low-velocity anomalies must be done with caution, and extra constraints from surface waves are needed (Cheng et al. 2015).

On the other hand, in some continental regions, such as in North America, owing to the recent dense USArray deployment, improved resolution is progressively achieved (e.g., Burdick et al. 2008; Obrebski et al. 2011; Sigloch and Mihalynuk 2013). Nevertheless, at the global scale, resolution from body wave tomography remains uneven, even when surface or core reflected teleseismic phases are added. Also, these tomographic models generally provide high-resolution information on *P* velocity, since S wave travel times are more difficult to measure accurately.

4.2.2.2 Surface Wave Tomography

Because their energy is concentrated near the surface along the source station great circle path, fundamental mode surface waves, in turn, allow the sampling of the upper mantle under oceans and continents alike. This leads to robust resolution of the long-wavelength component of lateral heterogeneity in shear velocity in the upper mantle at the global scale. However, because the sensitivity to structure decreases exponentially with depth, resolution from fundamental mode surface wave tomography is best in the first 300 km of the upper mantle. In order to improve resolution at larger depths, i.e., into the transition zone , it is necessary to include surface wave overtone data (e.g., Debayle and Ricard 2012; Schaefer and Lebedev 2015). These have similar group velocities, and hence, sophisticated approaches are required to separate and measure dispersion on individual overtone branches [see review by Romanowicz (2002)]. This presents a challenge for achieving comparable coverage to fundamental mode surface waves at the global scale.

This is why the recent global whole-mantle shear-velocity models that provide the best resolution in the transition zone (Kustowski et al. 2008; Ritsema et al. 2011) are based on a combination of different types of data which provide complementary sampling of the mantle: (1) fundamental mode surface waves and overtones, which provide resolution across the upper mantle; (2) for the lower mantle, body wave travel times, which generally include, in addition to first arriving S waves, surface reflected SS and core reflected ScS waves, sometimes complemented by core-propagating SKS travel time data. Some models, based on secondary travel time observables, also consider another type of data, normal mode "splitting functions", which provide constraints on the longest wavelength structure throughout the mantle (e.g., Ritsema et al. 2011).

4.2.2.3 Global Waveform Tomography Based on Asymptotic Methods

Since body and surface waves sample the earth differently , a powerful way to improve the sampling of the mantle is to combine them by exploiting the information contained in the entire seismogram (i.e., seismic waveforms). This idea was first introduced in global tomography by Woodhouse and Dziewonski (1984), where observed and synthetic seismograms were directly compared in the time domain. Introducing long-period seismic waveform tomography allowed these authors to include information from overtones in a simple way and thus to improve resolution in the transition zone. Synthetic seismograms were computed in a 3D earth using normal mode summation and the "path average" approximation (PAVA). A similar type approach has also been developed (Nolet 1990) and applied to upper mantle tomography at the continental (Van der Lee and Nolet 1997) and global scales (Lebedev and Van Der Hilst 2008; Schaeffer and Lebedev 2013).

In standard body wave tomography, the ensemble of body wave phases available through travel time measurements is largely limited. For example, the study of Kustowski et al. (2008) mentioned above was limited to measurements of SS,

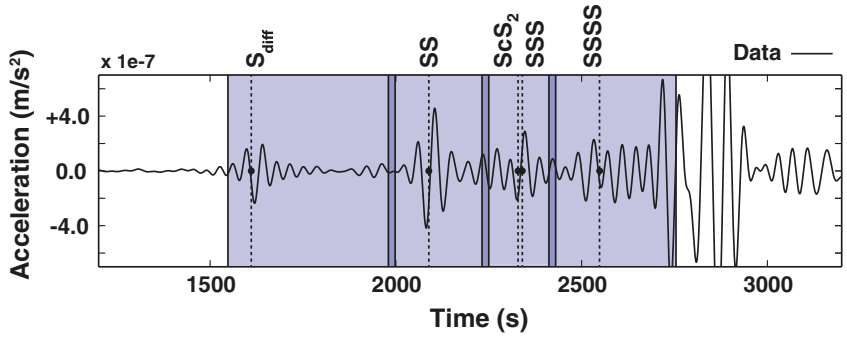

Fig. 4.2 Example of wave packet selection procedure for time-domain waveforms, as used in Mégnin and Romanowicz (2000) and following models from the Berkeley group, that are based on time-domain waveform inversion. *Shaded areas* indicate wave packets picked. Note in the third wave packet, the combination of two body wave phases (SSS, ScS2) that are not separable for travel time computation, but that sample very different parts of the mantle (courtesy of Scott French)

ScS, and SKS phases. A clear advantage of waveform tomography is that one can include body phases that cannot be separated in the time domain such as, for example, ScS_2 and SSS, as well as diffracted waves, whose propagation cannot be well described by ray theory (see Fig. 4.2).

However, when using body waveforms, the path average approximation (PAVA) may not be valid anymore. Indeed, the drawback of the PAVA is that it assumes that sensitivity of the waveforms is limited to the average 1D structure between the epicenter and the receiver, which is clearly inappropriate for body waves, whose sensitivity is concentrated along the ray path (Romanowicz 1987). In order to take into account the concentration of sensitivity along the ray path of body waves, across-branch coupling needs to be included (e.g., Li and Tanimoto 1993). Li and Romanowicz (1995) developed NACT (nonlinear asymptotic coupling theory), which introduced an additional term to PAVA that accounted for coupling across normal mode dispersion branches, bringing out the ray character of body waveforms [see Romanowicz et al. (2008) for details and a comparison of mode-based methods for modeling seismic waveforms]. This approach has been applied to the development of waveform-based global long-wavelength shear-velocity models since the mid-1990s (e.g., Li and Romanowicz 1996; Mégnin and Romanowicz 2000; Panning and Romanowicz 2006; Panning et al. 2010).

Comparing models obtained by different groups using different datasets and methodologies is one way to evaluate the robustness of the retrieved structure. The advantage of using full waveform tomography is that, by including a variety of phases that illuminate the mantle in different ways, the sampling is improved in ways that cannot be attained using only travel times of well isolated phases, largely because the distribution of earthquake sources and receivers is limited resulting in many redundant paths even as new data are added. Thus, at the very least, the same resolution can be achieved using considerably fewer source station paths. This is illustrated in Fig. 4.3 which shows a comparison of three recent global shear-velocity tomographic models at a depth of 100 km. Models (a and c)

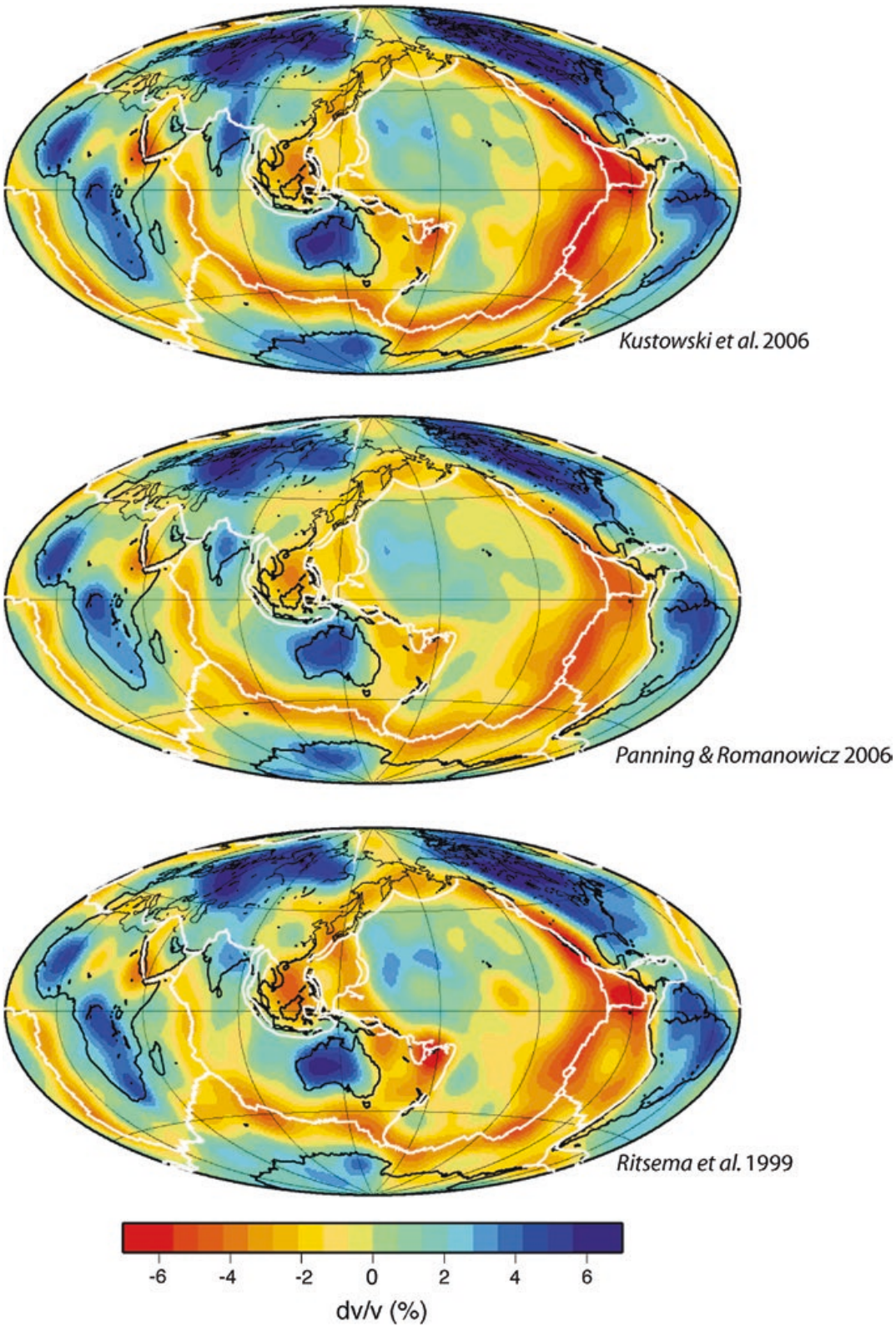
Kustowski et al. 2006
Panning & Romanowicz 2006
Ritsema et al. 1999
-6
-4
-2
0
2
4
6
dv/v (%)

Fig. 4.3 Comparison of maps of isotropic versus at a depth of 100 km from three whole-mantle tomographic models: **a** S362ANI (Kustowski et al. 2008); **b** SAW642AN (Panning and Romanowicz 2006); and **c** S20RTS (Ritsema et al. 1999). Model **a** was constructed using a combination on body wave travel times, surface wave dispersion, and long-period waveforms, albeit with the PAVA approximation; Model **c** was constructed using a combination of surface wave dispersion and body wave travel times. Both models used over 200,000 data. Model **b** was constructed using time-domain waveforms exclusively and the NACT theoretical framework, obtaining an equivalent resolution to the two other models, albeit with an order of magnitude fewer data (20,000 waveform packets)

were obtained using a conventional approach: Ritsema et al. (1999) used over 2M fundamental mode and overtone measurements combined with over 20,000 body wave travel time measurements to construct model S20RTS (a), while Kustowski et al. (2008) used several million dispersion measurements and about 150,000 body wave travel time measurements to construct model 362ANI. In contrast, Panning and Romanowicz (2006) used "only" 20,000 long-period time-domain seismograms (i.e., waveforms) and NACT to construct model SAW36ANI and were able to resolve the long-wavelength structure in the upper mantle just as well. With the ability to include increasingly shorter periods, i.e., constraints from phases that sample the mantle in yet other ways, as well as improving the accuracy with which the interactions of the wave-field with heterogeneity are computed, this opens the way to increased resolution in the future, as will be discussed in the next section. For now, beyond details of the datasets and theories used, Figs. 4.3 and 4.4 indicate that the level of agreement between global shear-velocity models is presently excellent up to at least degree 12 in a spherical harmonics expansion of the model, both in the upper and the lowermost mantle (e.g., Lekic et al. 2012).

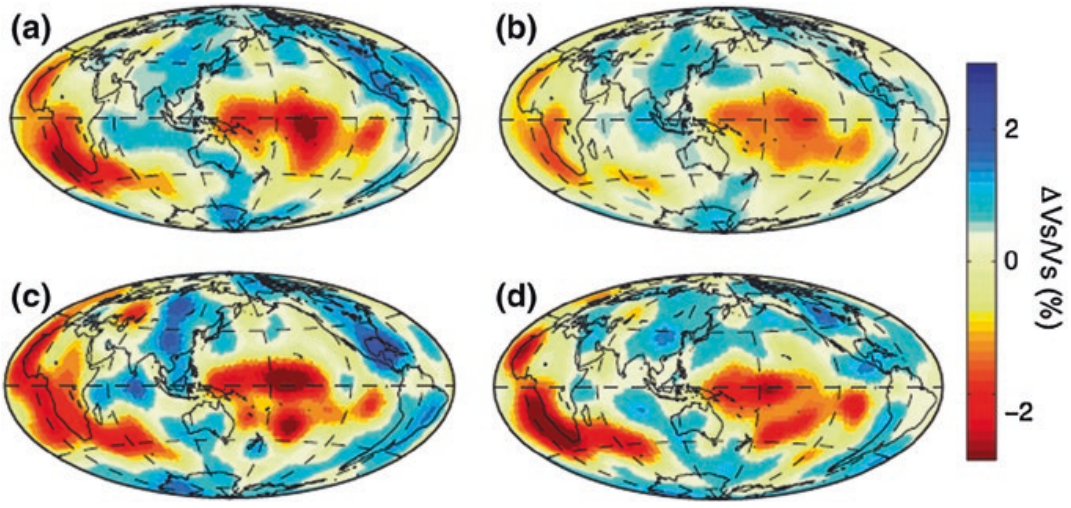

Fig. 4.4 Comparison of 4 recent shear-wave tomographic models at a depth of 2800 km. **a** Kustowski et al. (2008). **b** Ritsema et al. (2011). **c** Mégnin and Romanowicz (2000). **d** Houser et al. (2008). Model **c** was developed using only time-domain waveforms (about 20,000), while all other models are based on a combination of secondary observables (travel times of body waves and surface waves), except model **a** which includes long-period waveforms, albeit in the surface wave (PAVA) approximation. After Lekic et al. (2012)

4.2.2.4 Global Waveform Inversion Based on Direct Numerical Solvers

In the previous section, we have described how, in principle, full waveform tomography provides access to more of the information contained in seismograms than a collection of travel times of a limited number of seismic phases. As mentioned above, normal mode summation has provided a successful theoretical approach for computation of waveforms and led to several generations of whole-mantle shear-velocity models in the last 20 years. However, asymptotic normal mode perturbation theory (Li and Romanowicz 1995) is only valid for earth models for which the wavelength of the structure is large compared to that of the seismic waves considered (i.e., smooth models) and heterogeneity is weak (nominally, lateral variations of up to ∼10 %). Yet, in the earth's boundary layers, i.e., in the upper mantle and in the D″ region, there is ample evidence for the presence of stronger heterogeneity, whereas throughout the mantle, heterogeneity at many different scales may be present. First-order mode perturbation theory is not appropriate in this case, and more accurate numerical methods must be used. The challenge then is how to compute the synthetic seismograms in a 3D earth model without the weak heterogeneity approximation.

Finite difference methods are the traditional approach used for numerical calculation of seismograms (Kelly et al. 1976; Virieux 1986). In the 1990s, pseudo-spectral methods have also become a popular alternative and have been applied to regional (Carcione 1994) and global (Tessmer et al. 1992) problems. However, both finite difference and pseudo-spectral schemes perform poorly at representing surface waves. This issue can be addressed with the spectral element method (SEM) where the wave equation is solved on a mesh that is adapted to the free surface and to the main internal discontinuities of the model. The SEM was first introduced by Priolo et al. (1994) and Seriani and Priolo (1994) for wave-field calculation in 2D and later perfected by Komatitsch and Vilotte (1998), Komatitsch and Tromp (1999), and Komatitsch and Tromp (2002) for the 3D case. See Virieux and Operto (2009) for a review of numerical solvers in exploration geophysics.

Although these approaches started earlier in the exploration community than in global seismology, they are now reaching similar advance levels. Numerically computed seismograms automatically contain the full seismic wave-field, including all body and surface wave phases as well as scattered waves generated by lateral variations of the model Earth properties. The amount of exploitable information is thus significantly larger than in methods mentioned above. The accuracy of the numerical solutions and the exploitation of complete waveform information result in tomographic images that are both more realistic and better resolved (Fichtner et al. 2010). In seismology, the use of SEM has now been applied to tomographic inversions for crustal structure at the local scale (e.g., Tape et al. 2010) and upper mantle structure at regional scales (e.g., Fichtner et al. 2009, 2010; Rickers et al. 2013; Zhu et al. 2012; Zhu and Tromp 2013).

The forward numerical computation is generally combined with an "adjoint" formulation for the numerical computation of the kernels for inversion (Tromp et al. 2005; Fichtner et al. 2006) or, alternatively, with a "scattering integral

formalism" (e.g., Chen et al. 2007). In this context, Fichtner and Trampert (2011) showed how a local quadratic approximation of the misfit functional can be used for resolution analysis.

Here, we note that the inverse step is currently approached differently by different investigators. Following the nomenclature of the geophysical exploration community, the term FWI (full waveform inversion) is often used synonymously to "adjoint inversion," which relies, at each iteration, on the numerical computation of the gradient followed by a conjugate-gradient step. An alternative method, which has been used so far in SEM-based global waveform inversions, is to compute an approximate Hessian using mode-coupling theory in the current 3D model, followed by a Gauss–Newton (GN) inversion scheme. While it might be argued that the partial derivatives computed in this manner are more "approximate," the GN scheme is much faster converging (less than 10 iterations typically, compared to 30–40 or more) and can now take advantage of efficient methods for the assembly and inversion (ScalaPak) of large full matrices.

At the global scale, because the wave-field needs to be computed for a long time interval, in order to include all seismic phases of interest, the use of the SEM is particularly challenging computationally (Capdeville et al. 2005). Furthermore, computational time increases as the fourth power of frequency and limits the frequency range of waveforms to relatively long periods (typically longer than 40 or 50 s). The first global shear-velocity models developed using SEM (Lekić and Romanowicz 2011; French et al. 2013) are limited to the upper mantle due to the use of relatively long periods (longer than 60 s). In these models, the numerical computation of the forward step is restricted to the mantle and coupled with 1D mode computation in the core (CSEM, Capdeville et al. 2003). For the inverse step, kernels are computed using a mode-based approximation.

These modeling efforts have demonstrated the power of the SEM to sharpen tomographic images at the local, regional, and global scales and have led to the discovery of features previously not detected, such as the presence of low-velocity channels in the oceanic asthenosphere (e.g., French et al. 2013; Colli et al. 2013; Rickers et al. 2013). This is shown in Fig. 4.5 where model SEMum2 (French et al. 2013) is compared to other global shear-velocity models. SEMum2 more accurately recover both the depth and strength of the low-velocity minimum under ridges. It also shows stronger velocity minima in the low-velocity zone, a more continuous signature of fast velocities in subduction zones and stronger, clearly defined, low-velocity conduits under the Pacific Superswell, while confirming the robust long-wavelength structure imaged in previous studies, such as the progressive weakening and deepening of the oceanic low-velocity zone with overlying plate age. Of course, a more objective way to compare tomographic methods would be to conduct a blind test using numerically generated data, but this is beyond the scope of this study.

Because the frequency range of global inversions remains limited and because features smaller than the shortest wavelength cannot be mapped, this approach is not able, however, to resolve sharp discontinuities. The resulting tomographic images can therefore be seen as a smooth representation of the true Earth.

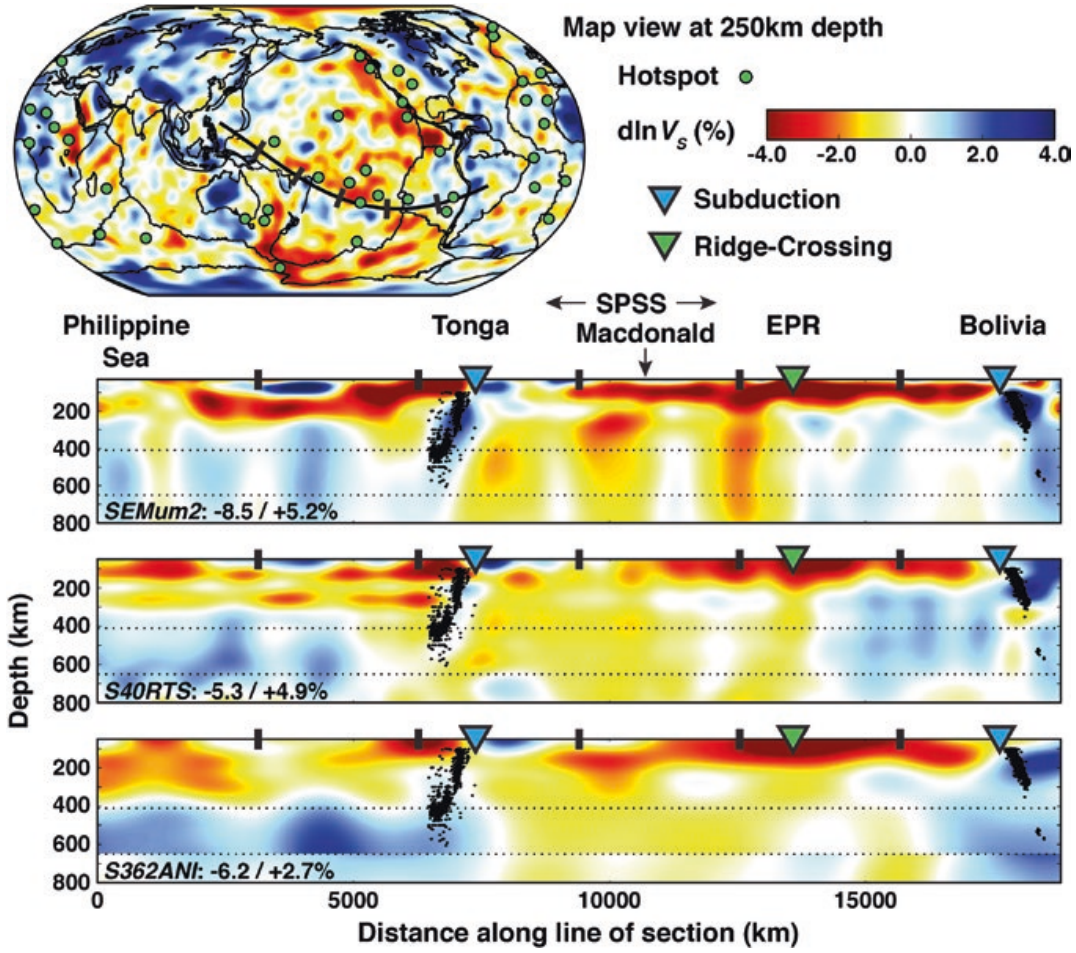

Fig. 4.5 Upper mantle depth cross sections across the Pacific superswell, comparing two recent global models obtained using classical approaches based on a combination of travel times, dispersion measurements, and approximate wave propagation theories (S362ANI, Kustowski et al. 2008; S40RTS, Ritsema et al. 2011) and a recent model constructed using waveforms and wavefield computations using SEM (SEMum2, French et al. 2013). While all three models agree in their long-wavelength structure in the transition zone, model SEMum2 shows more sharply delineated structures, both in subduction zones (highlighted by seismicity) and in the central Pacific, where the large low-velocity region is now resolved into two separate vertically oriented features. Model SEMum2 also exhibits stronger low-velocity minima in the uppermost mantle low-velocity zone. After French et al. (2013), courtesy of Scott French

However, they are not a simple spatial average of the true model, but rather an effective, apparent, or equivalent model that provides a similar long-wave data fit (Capdeville et al. 2010a, b). Hence, the geological interpretation of global tomographic models is limited, mainly due to two reasons:

1. The constructed images are smooth and do not contain discontinuities that are crucial to understand the structure and evolution of the earth.
2. The relations that link the true Earth to the effective (and unrealistic) earth that is seen by long-period waves are strongly nonlinear, and their inverse is highly non-unique. As a result, it is difficult to quantitatively interpret the level of imaged

anisotropy in tomographic models, as it may be the effect of either "real" local anisotropy or unmapped velocity gradients, or a combination of both.

4.3 Seismic Anisotropy

4.3.1 Observation of Anisotropy

It is well known that anisotropic structure is needed to predict a number of seismic observations such as:

1. Shear-wave splitting (or birefringence), the most unambiguous observation of anisotropy, particularly for SKS waves (Vinnik et al. 1989).
2. The Rayleigh–Love wave discrepancy. At global as well as at regional scale, the lithosphere appears faster to Love waves than to Rayleigh waves. It is impossible to simultaneously explain Rayleigh and Love wave dispersion by a simple isotropic model (Anderson 1961).
3. Azimuthal variation of the velocity of body waves. For example, Hess (1964) showed that the azimuthal dependence of Pn-velocities below oceans can be explained by anisotropy.

The goal here is not to provide a review of seismic anisotropy, but to address the issue of separating intrinsic and extrinsic anisotropy in apparent (observed) anisotropy. There are excellent review papers and books that have been written on anisotropy. For example, the theory of seismic wave propagation in anisotropic media has been described in Crampin (1981), Babuška and Cara (1991), and Chapman (2004). See also Maupin and Park (2014) for a review of observations of seismic anisotropy. Montagner (2014) gives a review of anisotropic tomography at the global scale. Montagner (1994); Montagner and Guillot (2002) give a review of geodynamic implications of observed anisotropy. Finally, a review of the significance of seismic anisotropy in exploration geophysics has been published by Helbig and Thomsen (2005).

Seismic waves are sensitive to the full elastic tensor (21 parameters), density, and attenuation. As seen above, it is not possible to resolve all 21 components of the anisotropic tensor at every location. Therefore, seismologists rely on simplified (yet reasonable) assumptions on the type of anisotropy expected in the earth's upper mantle, namely hexagonal symmetry. This type of anisotropy (commonly called transverse isotropy) is defined by the 5 Love parameters A, C, F, L, and N (love 1927), and two angles describing the tilt of the axis of symmetry (Montagner and Nataf 1988). In this manuscript, we will limit ourselves to the case of radial anisotropy, which corresponds to transverse isotropy with a vertical axis of symmetry and no azimuthal dependence.

It can be shown (Anderson 1961; Babuška and Cara 1991) that for such a vertically transversely isotropic (VTI) medium, long-period waveforms are primarily sensitive to the two parameters:

$$V_{SH} = \sqrt{\frac{N}{\rho}} \tag{4.1}$$

$$V_{SV} = \sqrt{\frac{L}{\rho}} \tag{4.2}$$

where ρ is density, V_{SV} is the velocity of vertically traveling S waves or horizontally traveling S waves with vertical polarization, and V_{SH} is the velocity of horizontally traveling S waves with horizontal polarization. The influence of other parameters A (related to V_{VH}), C (related to V_{PV}), and F can be large (Anderson and Dziewonski 1982) and is usually taken into account with petrological constraints (Montagner and Anderson 1989). That is, once V_{SH} and V_{SV} are constrained from long-period seismic waves, the rest of the elastic tensor and density is retrieved with empirical scaling laws (e.g., Montagner and Anderson 1989). Globally, SH waves propagate faster than SV waves in the upper mantle. The velocity difference is of about 4 % on average in the preliminary reference Earth model (PREM) of PREM Dziewonski and Anderson (1981) in the uppermost 220 km of the mantle.

Although early global radially anisotropic models were developed in terms of V_{SH} and V_{SV}, more recent models are parameterized in terms of an approximate Voigt average isotropic shear velocity (Montagner 2014) and radial anisotropy as expressed by the ξ parameter (e.g., Gung et al. 2003; Panning and Romanowicz 2006):

$$V_S = \frac{2V_{SV} + V_{SH}}{3} = \sqrt{\frac{2L + N}{3\rho}} \tag{4.3}$$

$$\xi = \frac{V_{SH}^2}{V_{SV}^2} = \frac{N}{L} \tag{4.4}$$

4.3.2 Anisotropy of Minerals: Intrinsic Anisotropy

Anisotropy can be produced by multiple physical processes at different spatial scales. It exists from the microscale (crystal scale) to the macroscale, where it can be observed by seismic waves that have wavelengths up to hundreds of kilometers. We name intrinsic anisotropy, the elastic anisotropy still present whatever the scale of investigation, down to the crystal scale. Most minerals in the earth's upper mantle are anisotropic. Olivine, the most abundant mineral in the upper mantle, displays a P wave anisotropy larger than 20 %. Other important constituents such as orthopyroxene or clinopyroxene are anisotropic as well (>10 %). Under finite strain accumulation, plastic deformation of these

minerals can result in a preferential orientation of their crystalline lattices. This process is usually referred to as LPO (lattice-preferred orientation) or CPO (crystalline-preferred orientation). This phenomenon is often considered as the origin of the observed large-scale seismic anisotropy in the upper mantle. With increasing the depth, most of minerals undergo a series of phase transformations. There is some tendency (though not systematic) that with increasing pressure, the crystallographic structure evolves toward a more closely packed, more isotropic structure, such as cubic structure. For example, olivine transforms into β-spinel and then γ-spinel in the upper transition zone (410–660 km of depth) and into perovskite and magnesiowustite in the lower mantle and possibly into post-perovskite in the lowermost mantle. Perovskite, post-perovskite $(Mg,Fe)SiO_3$, and the pure end-member of magnesiowustite MgO are still anisotropic. That could explain the observed anisotropy in some parts of the lower mantle and D″-layer.

Mantle rocks are assemblages of different minerals which are more or less anisotropic. The resulting amount of anisotropy is largely dependent on the composition of the aggregates. The relative orientations of crystallographic axes in the different minerals must not counteract in destroying the intrinsic anisotropy of each mineral. For example, the anisotropy of peridotites, mainly composed of olivine and orthopyroxene, is affected by the relative orientation of their crystallographic axes, but the resulting anisotropy is still larger than 10 %.

In order to observe anisotropy due to LPO at very large scale, several conditions must be fulfilled. The crystals must be able to re-orient in the presence of strain and the deformation due to mantle convection must be coherent over large scales to preserve long-wavelength anisotropy. These processes are well known for the upper mantle, and in oceanic plates, and anisotropy remains almost uniform on horizontal length scales in excess of 1000 km. The mechanisms of alignment are not so well known in the transition zone and in the lower mantle. In addition, a significant water content such as proposed by Bercovici and Karato (2003) in the transition zone can change the rheology of minerals, would make the deformation of the minerals easier, and change their preferential orientation. A complete discussion of these different mechanisms at different scales can be found in Mainprice (2007).

At slightly larger scale (but smaller than the seismic wavelength), a coherent distribution of fluid inclusions or cracks (Crampin and Booth 1985) can give rise to apparent anisotropy due to shape-preferred orientation (SPO). This kind of anisotropy related to stress field can be considered as the lower lilt of extrinsic anisotropy.

Anisotropic properties of rocks are closely related to their geological history and present configuration and reveal essential information about the earth's structure and dynamics (Crampin 1981). This justifies the great interest of geophysicists in all seismic phenomena which can be interpreted in the framework of anisotropy. However, the observation of large-scale anisotropy is also due to other effects such as unmapped velocity gradients.

4.3.3 Apparent Anisotropy Due to Small-Scale Inhomogeneities

It has been known for a long time in seismology and exploration geophysics that small-scale inhomogeneities can map into apparent anisotropy (Postma 1955; Backus 1962). The problem is very well described in the abstract by Levshin and Ratnikova (1984): "ln homogeneities in a real material may produce a seismic wave-field pattern qualitatively indistinguishable from one caused by anisotropy. However, the quantitative description of such a medium as an apparently anisotropic elastic solid may lead to geophysically invalid conclusions."

The scattering effect of small-scale heterogeneities on seismograms has been extensively studied in seismology (e.g., Aki 1982; Richards and Menke 1983; Park and Odom 1999; Ricard et al. 2014). As an example, Kennett and Nolet (1990) and Kennett (1995) demonstrated the validity of the great circle approximation when modeling long-period waveforms. However, despite all these studies, poor attention has been given to the theoretical relations between small-scale heterogeneities and equivalent anisotropy. By definition, an anisotropic material has physical properties which depend on direction, whereas a heterogeneous material has properties which depend on location. But the distinction between heterogeneity and anisotropy is a matter of the scale at which we analyze the medium of interest. Alternating layers of stiff and soft material will be seen at large scales as a homogeneous anisotropic material. At the origin of any anisotropy, there is a form for heterogeneity. In this way, the most basic form of anisotropy, related to the regular pattern made by atoms in crystals, can also be seen as some form of heterogeneity at the atomic scale (Maupin and Park 2014).

Although poorly studied theoretically, this phenomenon has been recognized in a number of studies. Maupin (2001) used a multiple-scattering scheme to model surface waves in 3D isotropic structures. She found that the apparent Love–Rayleigh discrepancy ($V_{SH} - V_{SV}$) varies linearly with the variance of isotropic S wave velocity anomalies. In the case of surface wave phase-velocity measurements done at small arrays, Bodin and Maupin (2008) showed that heterogeneities located close to an array can introduce significant biases which can be mistaken for anisotropy. For the lowest mantle, Komatitsch et al. (2010) numerically showed that isotropic velocity structure in D″ can explain the observed splitting of S_{diff}, traditionally interpreted as LPO intrinsic anisotropy due to mantle flow.

In the context of joint inversion of Love and Rayleigh waveforms, a number of studies acknowledged that the strong mapped anisotropy is difficult to reconcile with mineralogical models. This discrepancy may be explained in part by horizontal layering, or by the presence of strong lateral heterogeneities along the paths, which are simpler to explain by radial anisotropy (Montagner and Jobert 1988; Friederich and Huang 1996; Ekström and Dziewonski 1998; Debayle and Kennett 2000; Raykova and Nikolova 2003; Endrun et al. 2008; Bensen et al. 2009; Kawakatsu et al. 2009).

Bozdağ and Trampert (2008) showed that the major effect of incorrect crustal corrections in surface wave tomography is on mantle radial anisotropy. This is

because the lateral variation of Moho depth trade-offs with radial anisotropy [see also Montagner and Jobert (1988), Muyzert et al. (1999), Lebedev et al. (2009), Lekić et al. (2010), and Ferreira et al. (2010)].

Therefore, it is clear that both vertical and lateral isotropic heterogeneities can contribute to the observed radial anisotropy. The problem of separating intrinsic and apparent anisotropy is too complex in full generality. We can, however, examine a simple and illustrative problem. Following the recent work of Wang et al. (2013) and Fichtner et al. (2013a), we will place ourselves in the 1D radially symmetric case (VTI medium) and assume that apparent radial anisotropy is only due to vertical gradients, i.e., layering. Indeed, apart from the crust, the D' layer and around subducting slabs, to first order the earth is radially symmetric, with sharp horizontal seismic discontinuities separating different "layers" PREM (Dziewonski and Anderson 1981). In such a layered earth, vertical velocity gradients are much stronger than lateral ones and will significantly contribute to apparent anisotropy.

4.4 The Elastic Homogenization

We have seen that the limited resolution of long-wavelength seismic tomography only allows us to probe a smooth representation of the earth. However, this smooth equivalent Earth is not a simple spatial average of the true Earth, but the result of highly non-linear "upscaling" relations. In solid mechanics, these "upscaling" relations that link properties of a rapidly varying elastic medium to properties of the effective medium as seen by long waves have been the subject of extensive research (e.g., Hashin and Shtrikman 1963; Auriault and Sanchez-Palencia 1977; Bensoussan et al. 1978; Sanchez-Palencia 1980; Auriault et al. 1985; Murat and Tartar 1985; Sheng 1990; Allaire 1992, and many others).

In global seismology, upscaling schemes, also called elastic homogenization, have been recently developped for different kinds of settings (Capdeville and Marigo 2007; Capdeville et al. 2010a, b; Guillot et al. 2010; Capdeville et al. 2015). This class of algorithms enables to compute the effective properties of complex media, thus reducing the meshing complexity for the wave equation solver and hence the cost of computations. Elastic homogenization has been used to model complex crustal structures in full waveform inversions (Fichtner and Igel 2008; Lekić et al. 2010) and to combine results from different scales (Fichtner et al. 2013b).

4.4.1 The Backus Homogenization

Following the pioneering work by Thomson (1950), Postma (1955), and Anderson (1961), it was shown by Backus (1962) that a vertically transversely isotropic (VTI) medium is a "long-wave equivalent" to a smoothly varying medium of same

nature (i.e., transversely isotropic). For parameters concerning shear-wave velocities, the smooth equivalent medium is simply described by the arithmetic and harmonic spatial average of elastic parameters N and L:

$$\tilde{N} = \langle N \rangle \tag{4.5}$$

$$\tilde{L} = \langle 1/L \rangle^{-1} \tag{4.6}$$

where $\langle \cdot \rangle$ refers to a spatial average with length scale given by the shortest wavelength defining our "long-wave." In the rest of the manuscript, the symbol ~ will be used to describe long-wave equivalent parameters. Note that these two relations are analogous to computing the equivalent spring constant (or equivalent resistance) when multiple springs (or resistors) are mounted either in series or parallel. In simple words, a horizontally traveling wave V_{SH} will see a set of fine horizontal layers "in parallel" (5), whereas a vertically traveling wave V_{SV} will see them "in series" (6). The apparent density $\tilde{\rho}$ is also given by the arithmetic mean of the local density:

$$\tilde{\rho} = \langle \rho \rangle \tag{4.7}$$

In the case of a locally isotropic medium ($N = L$), i.e., with no intrinsic anisotropy, the homogeneous anisotropy is simply given by the ratio of arithmetic to harmonic mean:

$$\tilde{\xi} = \frac{\tilde{N}}{\tilde{L}} = \langle N \rangle \langle 1/N \rangle \tag{4.8}$$

It can be easily shown that the arithmetic mean is always greater than the harmonic mean, which results in having artificial anisotropy in (8) always greater than unity in the case of an underlying isotropic model. In the case where the underlying layered model contains anisotropy ($N \neq L$), the observed anisotropy is given by

$$\tilde{\xi} = \frac{\tilde{N}}{\tilde{L}} = \langle N \rangle \langle 1/L \rangle \tag{4.9}$$

Here, it is clear that when inverting waveforms with a minimum period of ~40 s (i.e., with minimum wavelength is 160 km) that sample a medium with velocity gradients occurring at much smaller scales, the observed apparent anisotropy $\tilde{\xi}$ is going to be different from the intrinsic anisotropy $\xi = N/L$. Therefore, as shown by Wang et al. (2013) and Fichtner et al. (2013a), interpreting the observed effective $\tilde{\xi}$ in terms of ξ may lead to misinterpretations.

4.4.2 The Residual Homogenization

In this study, the goal was to interpret smooth tomographic models in terms of their layered and hence more realistic equivalent. However, tomographic models

are not completely smooth and they are instead constructed as smooth anomalies around a discontinuous reference model. This is because the function linking the unknown model to the observed waveforms is linearized around a local point in the model space. This reference model often contains global discontinuities such as the Moho, or transition zone discontinuities at 410 and 660 km, which are fixed in the inversion and preserved in the model construction.

In the previous section, we have summarized an absolute homogenization for which no small scale is left in the effective medium. To account for the presence of a reference model, Capdeville et al. (2013), Capdeville and Cance (2015) recently described a modified homogenization, carried out with respect to a reference model, which we refer to as the residual homogenization. It allows us to homogenize only some interfaces of a discontinuous medium, while keeping the others intact.

Let us define the reference earth model by its density and elastic properties: (ρ_{ref}, A_{ref}, C_{ref}, F_{ref}, L_{ref}, and N_{ref}). Capdeville et al. (2013) showed that an equivalent model to the layered (A, C, F, L, and N) medium can be constructed with simple algebraic relations. For elastic parameters related to shear-wave velocities, we have:

$$\tilde{N} = N_{\text{ref}} + \langle N - N_{\text{ref}} \rangle \tag{4.10}$$

$$\frac{1}{\tilde{L}} = \frac{1}{L_{\text{ref}}} + \left\langle \frac{1}{L} - \frac{1}{L_{\text{ref}}} \right\rangle \tag{4.11}$$

Note that no particular assumption on the reference model is made, which can contain any wavelengths, and can be discontinuous. Furthermore, there is no linearity assumption, and this results holds for large differences between the reference and the layered model.

We show in Fig. 4.6 an example of residual homogenization. The layered VTI Model is shown in red with layers either isotropic ($\xi = 1$) or anisotropic ($\xi \neq 1$). A smooth equivalent model (for long waves of minimum wavelength of 100 km) that preserves the small scales of the reference model is shown in blue. The homogenization is done on the difference between the layered model in red and the reference model in thick light blue. After homogenization, we lose information about both the number and locations of discontinuities which are not in the reference model, as well as the location and level of intrinsic anisotropy.

It can be verified numerically that waveforms computed in the residual effective model and in the true layered model are identical when filtered with minimum period of 25 s (which corresponds to a minimum wavelength of 100 km). Figure 4.7 shows an example of seismogram computed by normal mode summation (Gilbert and Dziewonski 1975) in the residual effective model and compared with the solution computed in the true layered model. In both cases, the reference and homogenized traces show an excellent agreement.

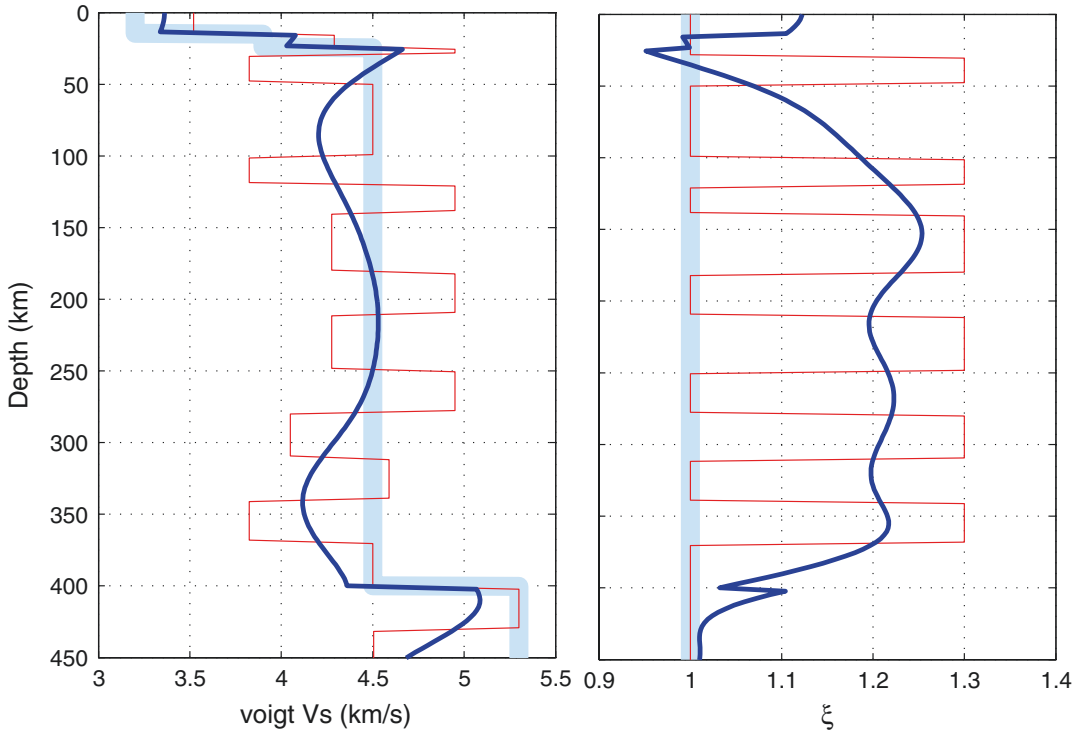

Fig. 4.6 Example of residual homogenization. *Left* Voigt average shear-wave velocity. *Right* Radial anisotropy. The layered model in *red* is homogenized around a reference model in light *blue*. The homogenized model is plotted in *blue*

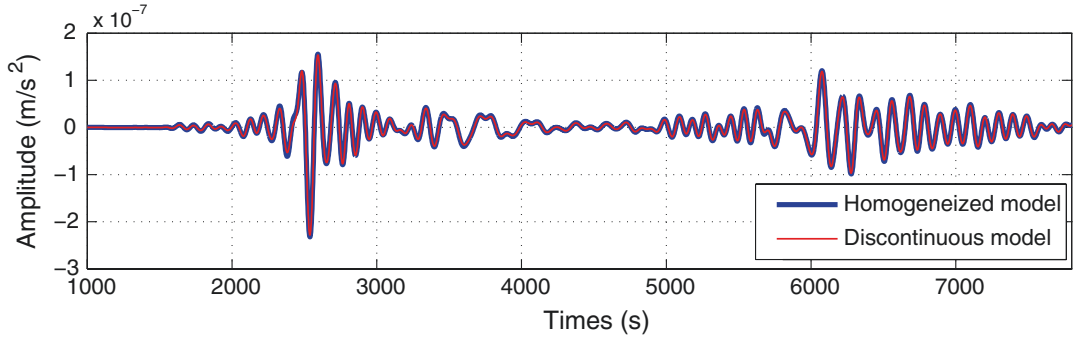

Fig. 4.7 Waveforms computed for the layered and homogenized models in Fig. 4.6. This is the radial component for an event with epicentral distance 82° and depth 150 km. The computation was done by normal mode summation (Gilbert and Dziewonski 1975)

4.4.3 An Approximation of "the Tomographic Operator"

Global full waveform tomography is always carried out with frequency band limited data. Intuitively, it makes sense to assume that such inversions can retrieve, at best, what is "seen" by the wave-field, i.e., an homogenized equivalent, and not the real medium.

Although it is difficult to mathematically prove this conjecture in general, Capdeville et al. (2013) numerically showed with synthetic examples that this is indeed the case for VTI media. That is, the inverted medium coincides with the residual homogenized version of the target model. Given a radially symmetric Earth and given enough stations and earthquakes, an inversion of full waveforms carried out around a reference model will therefore produce the residual homogeneous model defined above.

In this way, for any given layered model, one is able to predict with simple nonlinear algebraic smoothing operations what an inversion will find, without actually running the inversion. Therefore, we can view the residual homogenization as a first-order approximation of the "tomographic operator."

In practice, several practical issues complicate the situation: The real inversions are damped, producing unknown uncertainties in the recovered model, which can potentially bias our results. Furthermore, as seen above, ray coverage is not perfect and tomographic schemes may actually recover less than the effective medium.

4.5 Downscaling Smooth Models: The Inverse Homogenization

As we have seen, a tomographic inversion of long-period waves can only retrieve at best a homogenized model (and less in the case of an incomplete data coverage). Homogenization can lead to non-trivial and misleading effects that can make the interpretation difficult. We propose to treat the interpretation of tomographic images in terms of geological structures (discontinuities in our layered case) as a separate inverse problem, allowing to include a priori information and higher frequency data.

We call this inverse problem the inverse homogenization: For a given smooth 1D profile extracted from a tomographic model, what are the possible fine-scale (i.e., layered) models that are equivalent to this smooth 1D profile? Since the upscaling relations are based on nonlinear smoothing operators, it is not trivial to invert them to derive the true Earth from tomographic images, i.e., from its residual equivalent. In this section, we show that, although there is an infinite number of layered models that are equivalent to the smooth model in blue (Fig. 4.6), these models share common features, and Bayesian statistics can be used to constrain this ensemble of possible models. Furthermore, higher frequency data sensitive to discontinuities in radially symmetric models, such as receiver functions, can be used to constrain the location of horizontal discontinuities and reduce the space of possible earths.

4.5.1 Major Assumptions

Given the simple machinery presented in previous sections, there are obvious limitations to the proposed procedure. Let us here acknowledge a few of them.

1. We will assume here that long-period waves are only sensitive to the elastic parameters N and L (i.e., V_{SH} and V_{SV}). However, in a VTI medium, long-period seismograms, and hence the observed radial anisotropy, are also sensitive to the 3 other Love parameters (i.e., A, C, and F). Fichtner et al. (2013a) recently showed that P wave anisotropy is also important to distinguish between intrinsic and extrinsic anisotropy. Here, P wave anisotropy will be ignored.
2. Here, we restrict ourselves to transverse isotropy with a vertical axis of symmetry. Although this simple parameterization in terms of radial anisotropy is widely used in global seismology, it clearly represents an over-simplification, adopted for convenience in calculation. This is because the separation of intrinsic and apparent anisotropy can be studied analytically. The Earth is certainly not transversely isotropic, and there are indisputable proofs of azimuthal anisotropy. Azimuthal anisotropy might map into radial anisotropy in global models. These effects could be analyzed using the 3D version of non-periodic homogenization (Capdeville et al. 2010a, b).
3. We assume that 1D vertical profiles extracted from 3D tomographic models are the true Earth that has been homogenized with Backus relations. However, the smoothing operator applied to the true Earth during an inversion, namely the "tomographic operator," is determined by an ensemble of factors such as poor data sampling, the regularization and parameterization imposed, the level of data noise, the approximations made on the forward theory, and limited frequency band. It is very difficult to estimate how these averaging processes are applied to the true Earth during a tomographic inversion. What we assume here is that all these effects are negligible compared to the last one (limited frequency band), for which the smoothing operator is simply given by elastic homogenization. This only holds if data sampling is perfect, if no strong regularization has been artificially applied, and if the forward theory is perfect. Therefore, it is going to be most true in the case of full waveform inversion, and full waveform tomographic models are the most adequate for such a procedure. However, it is clear that other types of observations could be used as any tomographic method unavoidably produces apparent anisotropic long-wavelength equivalents. For example, our proposed procedure could be used to describe the ensemble of discontinuous models that fit a set of dispersion curves as in Khan et al. (2011).

4.5.2 Bayesian Inference

Using the notation commonly employed in geophysical inversion, the problem consists in finding a rapidly varying model m, such that its homogenized equivalent profile $g(m)$ is "close" to a given observed smooth model d. Here, the forward function g is the residual homogenization procedure in (4.10) and (4.11). Since the long-period waveforms are sensitive to smooth variations of the

"Backus parameters" (Capdeville et al. 2013), the observed tomographic profile is parameterized as $d = \left[\tilde{N}, 1/\tilde{L}\right]$.

This takes the form of a highly nonlinear inverse problem, and a standard linearized inversion approach based on derivatives is not adequate since the solution would strongly depend on the initial guess. Furthermore, the problem is clearly under-determined and the solution non-unique, and hence, it does not make sense to look for a single best fitting model that will minimize a misfit measure $\|d - g(m)\|$. For example, one can expect strong correlations and trade-offs between unknown parameters as homogeneous anisotropy can be either explained by discontinuities or intrinsic anisotropy. An alternative approach is to embrace the non-uniqueness directly and employ an inference process based on parameter space sampling. Instead of seeking a best model within an optimization framework, one seeks an ensemble of solutions and derives properties of that ensemble for inspection. Here, we use a Bayesian approach and tackle the problem probabilistically (Box and Tiao 1973; Sivia 1996; Tarantola 2005). We sample a posterior probability distribution $p(m|d)$, which describe the probability of having a discontinuous model m given an observed tomographic homogeneous profile d.

An important issue is the degree of freedom in the layered model. Since the inverse homogenization is a downscaling procedure, the layered model may be more complex (i.e., described with more parameters) than its smooth equivalent. As discussed above, the smooth model may be equivalent to either isotropic models with a large number of spatial parameters (layers), or anisotropic models described with more than one parameter per layer. This raises the question of the parameterization of m. How many layers should we impose on m? Should the existence (or not) of anisotropy be a free parameter? If yes, how many isotropic and anisotropic layers?

We propose to rely on Occam's razor, or the principle of parsimony, which states that simple models with the least number of parameters should be preferred (Domingos 1999). The razor states that one should favor simpler models until simplicity can be traded for greater explanatory power. Although we acknowledge that the definition of "simplicity" is rather subjective, in our problem, we will be giving higher probability to layered models described with fewer parameters.

We impose on m to be described with constant velocity layers separated by infinite gradients. As shown in Fig. 4.8 , we use a transdimensional parameterization, where the number of layers, as well as the number of parameters per layer, is free variables, i.e., unknown parameters (Sambridge et al. 2013). In this way, the number of layers will be unknown in the inversion, as well as the number of parameters in each layer: 1 for isotropic layers (V_S) and 2 for anisotropic layers (V_S and ξ). The goal here is not to describe the algorithm and its implementation in detail, but instead to give the reader a general description of the procedure, and show how it can be used to distinguish between intrinsic and extrinsic anisotropy. For a details on the algorithm, we refer the reader to Bodin et al. (2012b).

Bayes' theorem (1763) is used to combine prior information on the model with the observed data to give the posterior probability density function:

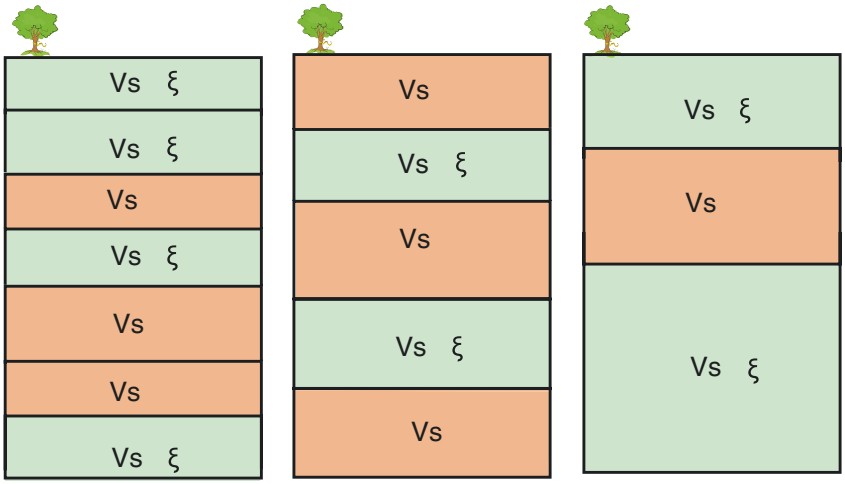

Fig. 4.8 Adaptive parameterization used for the inverse homogenization. The number of layers as well as the number of parameter in each layer (one for isotropic layers and two for anisotropic layers) is unknown in the inversion. This is illustrated here with three different models with different parameterizations. The parameterization is itself an unknown to be inverted for during the inversion scheme. Of course, data can always be better fitted as one includes more parameters in the model, but within a Bayesian formulation, preference will be given to simple models that explain observations with the least number of model parameters

$$\text{posterior} \propto \text{likelihood} \times \text{prior} \tag{4.12}$$

$$p(m|d) \propto p(d)|m)p(m) \tag{4.13}$$

$p(m)$ is the a priori probability density of m, that is, what we (think we) know about the model m before considering d. Here, we use poorly informative uniform prior distributions, and let model parameters vary over a large range of possible values.

The likelihood function $p(d|m)$ quantifies how equivalent a given discontinuous model is to our observed smooth profile d. The form of this probability density function is given by what we think about uncertainties on d. In our case, the form of the error statistics for a tomographic profile must be assumed to formulate $p(d|m)$. A problem with tomographic images is that they are obtained with linearized and regularised inversions, which biases uncertainty estimates. Therefore, we adopt a common and conservative choice (supported by the central limit theorem) and assume Gaussian-distributed errors. Since the data vector d is smooth, its associated errors must be correlated, and the fit to observations, $\Phi(m)$, is no longer defined as a simple 'least-square' measure but is the Mahalanobis distance between observed, d, and estimated, $g(m)$, smooth profiles:

$$\Phi(m) = (g(m) - d)^T C_e^{-1} (g(m) - d) \tag{4.14}$$

where C_e represents the covariance matrix of errors in d. In contrast to the Euclidean distance, this measure takes in account the correlation between data

(equality being obtained where C_e is diagonal). Note that there is no user-defined regularization terms in (4.14) such as damping or smoothing constraints. This misfit function only depends on the observed data.

The general expression for the likelihood probability distribution is hence:

$$p(d|m) = \frac{1}{\sqrt{(2\pi)^n |C_e|}} \times \exp\left\{ \frac{-\Phi(m)}{2} \right\}. \tag{4.15}$$

This is combined with the prior distribution to construct the posterior probability density function, which is thus defined in a space of variable dimension (transdimensional).

4.5.3 *Sampling a Transdimensional Probability Density Function*

Since the problem is transdimensional and nonlinear, there is no analytical formulation for the posterior probability density function, and instead we approximate it with a parameter search sampling algorithm (Monte Carlo). That is, we evaluate the posterior at a large number of locations in the model space. We use the reversible jump Markov chain Monte Carlo (rj-McMC) algorithm (Geyer and Møller 1994; Green 1995, 2003), which is a generalization of the well-known Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970) to variable dimension models. The solution is represented by an ensemble of 1D models with variable number of layers and thicknesses, which are statistically distributed according to the posterior distribution. For a review of transdimensional Markov chains, see Sisson (2005). For examples of applications in the Earth sciences, see Malinverno (2002), Dettmer et al. (2010), Bodin et al. (2012a), Ray and Key (2012), Iaffaldano et al. (2012, 2013), Young et al. (2013), Tkalčić et al. (2013), Pilia et al. (2015) and Choblet et al. (2014).

In order to illustrate the power of the proposed Bayesian scheme, we applied it to the synthetic homogenized profile shown in Fig. 4.6, polluted with some Gaussian random correlated (i.e., smooth) noise. The solution is a large ensemble of models parameterized as in Fig. 4.8, for which the statistical distribution approximates the posterior probability distribution. As will be shown below, there are a number of ways to look at this ensemble of models. Here, in Fig. 4.9, we simply plot the 2D marginal distribution on the number of layers and number of anisotropic layers. This allows us to quantify the trade-off between anisotropy and heterogeneity. The distribution is clearly bimodal, meaning that the smooth equivalent profile can either be explained by many isotropic layers or a few anisotropic ones. From this, it is clear that we have not been able to distinguish between real and artificial anisotropy. However, we are able (given a layered parameterization) to quantify probabilistically the non-uniqueness of the problem.

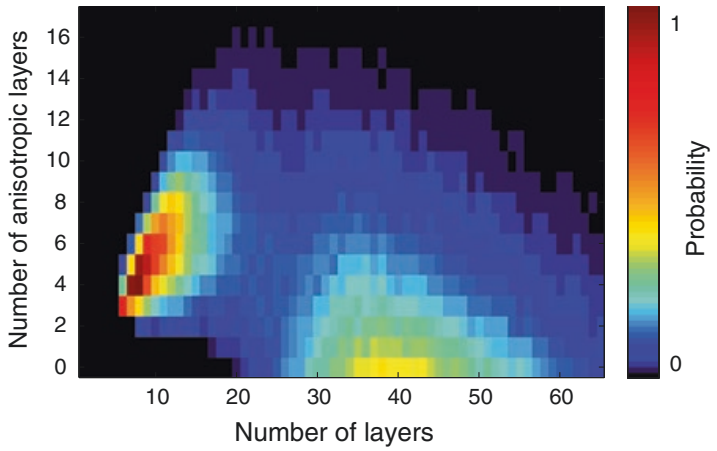

Fig. 4.9 Posterior probability distribution for the number of layers and number of anisotropic layers. This 2D marginal distribution allows us to quantify the trade-off between heterogeneity and anisotropy. Indeed, the smooth model in Fig. 4.6 can be either be explained with a large number of isotropic layers or a few anisotropic layers

This trade-off may be "broken" by adding independent constraints from other disciplines such as geology, mineral physics, or geodynamics. Here, we will show how higher frequency seismic data can bring information on the number and locations of discontinuities and hence enable us to investigate the nature of radial anisotropy in tomographic models.

4.6 Incorporating Discontinuities with Body Waves—Application to the North American Craton

A smooth equivalent profile brings little information about location of discontinuities, and extra information from higher frequency data is needed. Here, we show in a real case how adding independent constraints from converted P to S phases can help locating interfaces. Again, here, we place ourselves in the simplest case and assume horizontal layering when modeling converted phases. We acknowledge that dipping interfaces, or a tilted axis of anisotropy would produce apparent azimuthal anisotropy. Accounting for these effects is the subject of current work. We construct a 1D probabilistic seismic profile under northwest Canada, by combining in a joint Bayesian inversion a full waveform tomographic profile (SEMum2, French et al. 2013) with receiver functions. The goal here is to incorporate horizontal lithospheric discontinuities into a smooth image of the upper mantle and thus investigate the structure and history of the North American craton.

Archean cratons form the core of many of Earth's continents. By virtue of their longevity, they offer important clues about plate tectonic processes during early geological times. A question of particular interest is the mechanisms involved in cratonic assembly. The Slave province is one of the oldest Archaen cratons on Earth. Seismology has provided detailed information about the crust and upper mantle structure from different studies, such as reflection profiling (e.g., Cook et al. 1999), receiver function analysis (e.g., Bostock 1998), surface wave tomography (e.g., Van Der Lee and Frederiksen 2005), or regional full waveform (Yuan and Romanowicz 2010).

Recent studies (Yuan et al. 2006; Abt et al. 2010) have detected a structural boundary under the Slave craton at depths too shallow to be consistent with the lithosphere–asthenosphere boundary. Yuan and Romanowicz (2010) showed that this mid-lithospoheric discontinuity (MLD) may coincide with a change in the direction of azimuthal anisotropy and thus revealed the presence of two distinct lithospheric layers throughout the craton: a top layer chemically depleted above a thermal conductive root. On the other hand, Chen et al. (2009)showed that this seismic discontinuity as seen by receiver functions overlapped with a positive conductivity anomaly and interpreted it as the top of an archean subducted slab.

This type of fine structure within the lithosphere is not resolved in global tomographic models such as SEMum2 and hence may be mapped into radial anisotropy. Here, we will explore whether lithospheric layering as seen by scattered body waves (receiver functions) is compatible with the radial anisotropy imaged from global tomography.

4.6.1 Long-Period Information: A Smooth Tomographic Profile

We used the global model recently constructed by the Berkeley group: SEMum2 (Lekić and Romanowicz 2011; French et al. 2013). This model is the first global model where the synthetic waveforms are accurately computed in a 3D Earth with the spectral element method. Sensitivity kernels are calculated approximately using nonlinear asymptotic coupling theory (NACT: Li and Romanowicz (1995)). The database employed consists of long-period ($60 < T < 400$ s) three-component waveforms of 203 well-distributed global earthquakes ($6.0 < M_w < 6.9$), as well as global group-velocity dispersion maps at $25 < T < 150$ s.

Compared to other global shear-velocity models, the amplitudes of velocity anomalies are stronger in SEMum2, with stronger velocity minima in the low-velocity zone (asthenosphere) and a more continuous signature of fast velocities in subduction zones.

Here, we extract a 1D profile (Fig. 4.10) under station YKW3, located in the southern Slave craton, northwest Canada. As seen in Fig. 4.10, the crustal structure in SEMum2 is replaced with a single, smooth equivalent anisotropic layer, valid for modeling long-period waves. Note also that the high amplitude of radial anisotropy below the crust may be due to unmapped layering at these depths.

4.6.2 Short-Period Information: Teleseismic Converted Phases

In order to bring short-wavelength information to the tomographic profile, we analyzed waveforms for first P arrivals on teleseismic earthquake records at the

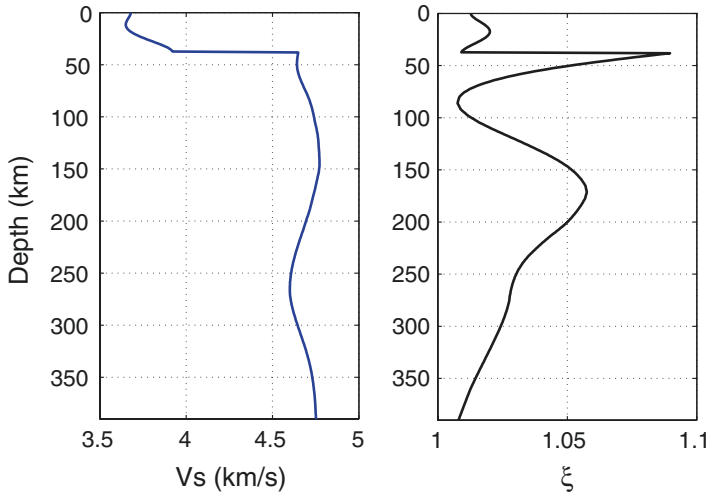

Fig. 4.10 Tomographic profile under station YKW3 for model SEMum2

broadband station YKW3 of the Yellowknife seismic array. The station was installed in late 1989 and has collected a large amount of data. Receiver function analysis consists of deconvolving the vertical from the horizontal component of seismograms (Vinnik 1977; Burdick and Langston 1977; Langston 1979). In this way, the influence of source and distant path effects are eliminated, and hence, one can enhance conversions from P to S generated at boundaries beneath the recording site. This is a widely used technique in seismology, with tens of papers published each year (e.g., Ford et al. 2010; Hopper et al. 2014). For a recent and comprehensive review, see Bostock (2014).

Algorithms for inversion of receiver functions are usually based on optimization procedures, where a misfit function is minimized. Traditionally, this misfit function is constructed by comparing the observed receiver function with a receiver function predicted for some Earth model m:

$$\Phi(m) = \left\| \frac{H(t)}{V(t)} - \frac{h(t,m)}{v(t,m)} \right\|^2 \tag{4.16}$$

where $V(t)$ is the vertical and $H(t)$ the horizontal (radial) component of the observed seismogram, and where $v(t,m)$ and $h(t,m)$ are predicted structure response functions for the unknown Earth model m. The fraction refers to a deconvolution (or spectral division).

A well-known problem is that the deconvolution is an unstable numerical procedure that needs to be damped, which results in a difficulty to correctly account for uncertainties. Therefore, for Bayesian analysis, we choose an alternative misfit function based on a simple cross-product that avoids deconvolution (Bodin et al. 2014):

$$\Phi(m) = \|H(t) * v(t,m) - V(t) * h(t,m)\|^2 \tag{4.17}$$

This misfit function is equivalent to the distance between the observed and predicted receiver functions in (4.16). Since discrete convolution in time is

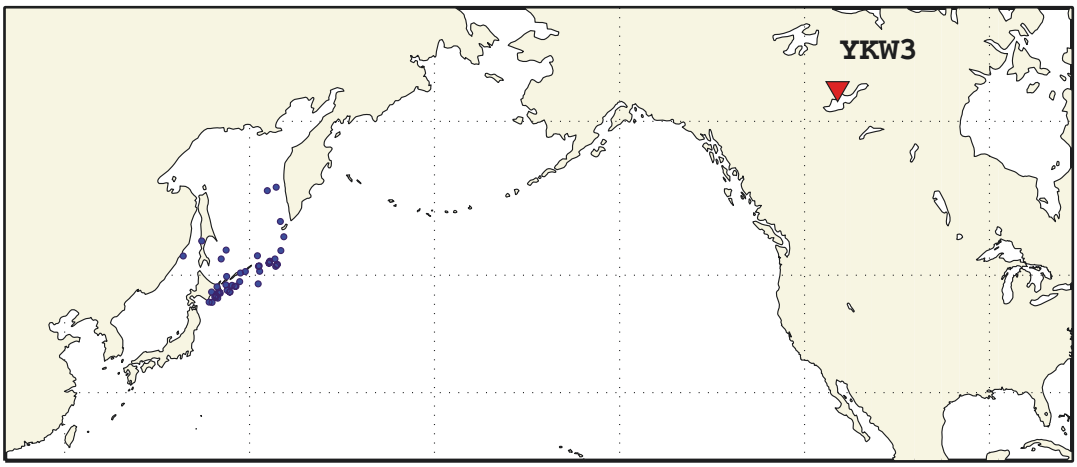

Fig. 4.11 Station YKW3 with the set of events used for receiver function analysis

a simple summation and since seismograms can be seen as corrupted by random errors, each sample of the signal obtained after discrete convolution is then a sum of random variables, whose statistics are straightforward to calculate with algebra of random variables. This is not the case with deconvolution schemes.

Assuming that $V(t)$ and $H(t)$ contain independent and normally distributed random errors with standard deviation σ, a likelihood probability function can be constructed:

$$p(d_{\mathrm{RF}}|m) = \frac{1}{\sqrt{(2\pi\sigma^2)^n}} \times \exp\left(\frac{-\Phi(m)}{2\sigma^2}\right) \tag{4.18}$$

The observed vertical $V(t)$ and horizontal $H(t)$ waveforms needed for inversion were obtained by simply stacking a number of events measured for a narrow range of backazimuths and epicentral distances (see Figs. 4.11 and 4.12). Influence of the receiver structure is common to all records and is enhanced by summation (Shearer 1991; Kind et al. 2012). We refer to Bodin et al. (2014) for details of the procedure.

This likelihood function thus defined for receiver functions $p(d_{\mathrm{RF}}|m)$ can be combined with the likelihood function defined above for the Inverse homogenization problem $p(d_{\mathrm{tomo}}|m)$ for joint inversion of short- and long-wavelength information. Since the observations given by the tomographic model d_{tomo} are independent of the receiver function observations d_{RF}, the complete posterior probability function is then defined as follows:

$$p(m|d_{\mathrm{tomo}}, d_{\mathrm{RF}}) \propto p(d_{\mathrm{tomo}}|m) \times p(d_{\mathrm{RF}}|m) \times p(m) \tag{4.19}$$

and can be sampled with the reversible jump algorithm described above.

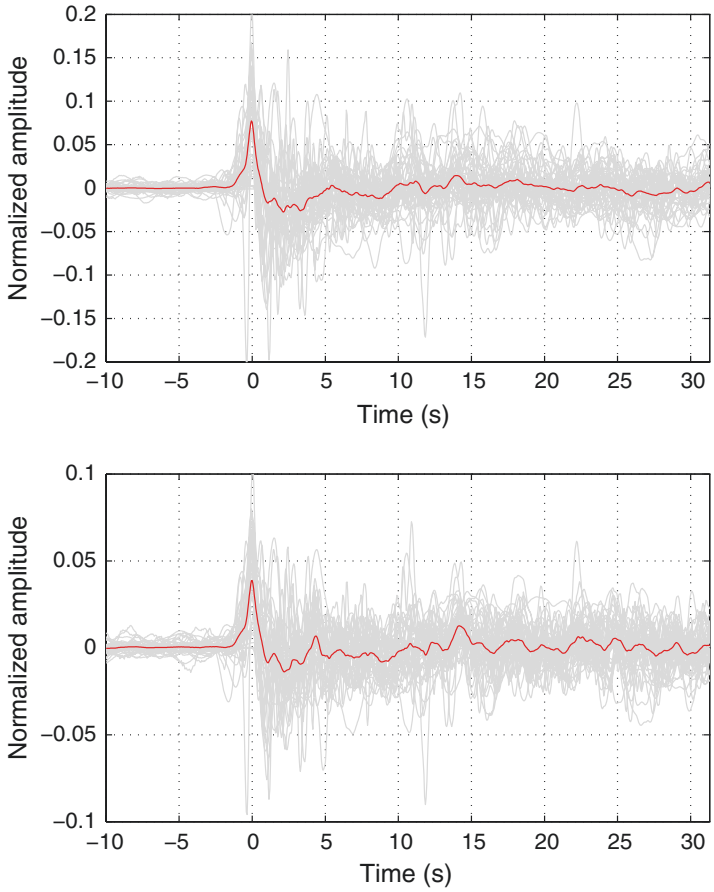

Fig. 4.12 Stack of P arrivals for receiver function analysis for event is shown in Fig. 4.11. A total of 44 events were used with backazimuths between 290° and 320° and with ray parameters between 0.04 and 0.045 s km^{-1}. Seismograms were cut for the same time window, normalized to equal energy, and rotated to radial and tangential components

4.6.3 Results

Transdimensional inversion was carried out allowing between 2 and 60 layers. As noted above, each layer is either described by one or two parameters. An a priori constrain for minimum and maximum velocity value and anisotropy in each layer was applied. The algorithm was implemented for parallel computers, providing a thorough search of the model space, with an ensemble solution made of about 10^6 different Earth models. The posterior distribution is approximated from the distribution (i.e., the histogram) of the ensemble of models in the solution (Fig. 4.13). The solution is thus given by an ensemble of 1D models with variable number of layers, thicknesses, and elastic parameters. In order to visualize the final ensemble, the collected models can be projected into a number of physical spaces that are used for interpretation.

For example, Fig. 4.13a shows the marginal distribution for S wave velocities as a function of depth, simply constructed from the density plot of the ensemble of models in the solution. Here, a number of expected lithospheric discontinuities have been imaged, such as the mid-lithospheric discontinuity at 90 km and a sharp lithosphere–asthenosphere boundary at 200 km. But interestingly, we also note a sharp positive velocity change at 150 km, which establishes the base of an

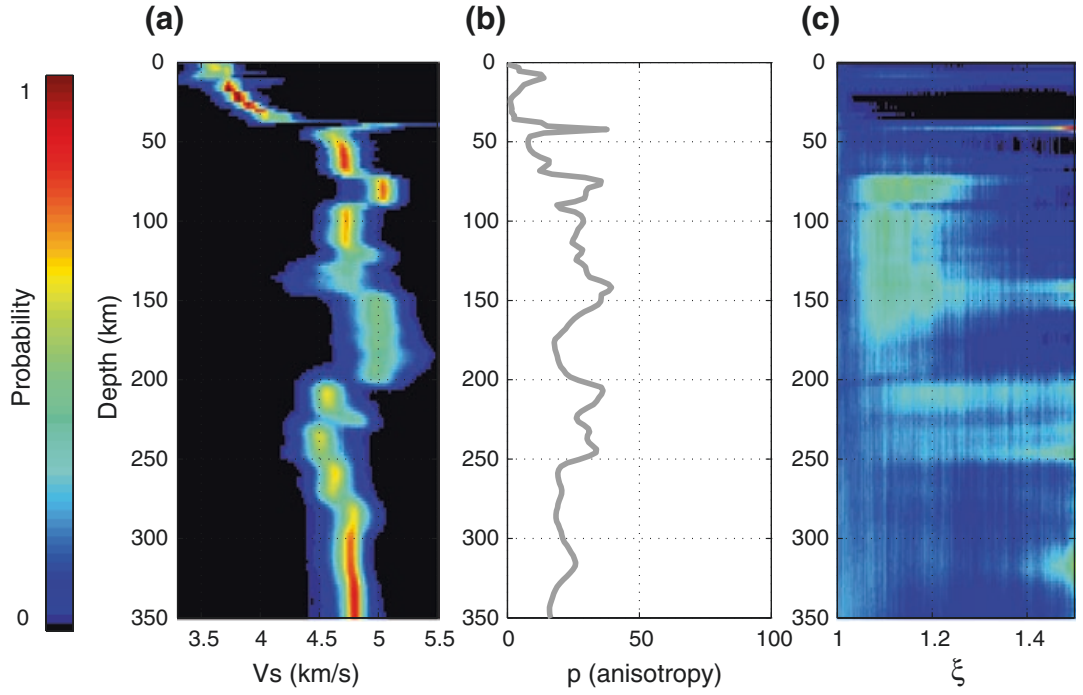

Fig. 4.13 Joint inversion of converted body waves in Fig. 4.12 and of the tomographic model in Fig. 4.10. *Left* Probability distribution for V_s. *Middle* Probability of having an anisotropic layer. *Right* Probability for ξ

intra-lithospheric low-velocity zone between 90 and 150 km. This low-velocity zone is clearly visible in Chen et al. (2007) who inverted Rayleigh wave phase-velocity observations (20 s $< T <$ 142 s) measured by the Yellowknife array. This feature is also observed in the regional full waveform tomographic model by Yuan and Romanowicz (2010). This low-velocity zone can be interpreted as a piece of archean subducted slab, stacked vertically over another archean block. For a discussion about upper mantel seismic discontibuities, see Schmerr (2015).

Since here we are interested in the relative contribution of layering to the observed anisotropy in SEMum2, we can look at the probability at each depth to have intrinsic anisotropy (i.e., $\xi \neq 1$). For each model, each layer is either isotropic or anisotropic. Therefore, at each depth, one can count the ratio of isotropic to anisotropic layers in the ensemble of models. This is shown in Fig. 4.13b. When no information is brought by the data (prior distribution), the probability to have anisotropy is 50 %. However here, the probability is lower and around 20 % across much of the depth profile. This implies that intrinsic anisotropy is not required to fit the smooth tomographic profile. In other words, the discontinuities required to fit the converted body waves may be enough to explain the anisotropy in the tomographic model. We view this as an important result, indicating that radial anisotropy (at least under cratons) should not be directly interpreted in tomographic models.

The third panel in Fig. 4.13c shows the probability distribution for ξ at each depth. Note that, for a given depth, around 80 % of models are isotropic with $\xi = 1$, and this distribution only represents the level of anisotropy in the 20 % remaining models.

Note that these are only preliminary results. This study is only a proof of concept, or study of feasibility, which opens a range of potential applications. Let us acknowledge a number limitations:

1. Here, we assume that smooth tomographic profiles are the true Earth that has been homogenized. Although this has been numerically demonstrated on 1D synthetic tests for VTI models (Capdeville et al. 2013), it may not always be true. For example, the effect of poor data sampling and hence model regularization may act as a smoothing operator not accounted for here.
2. Little is known about uncertainties in tomographic models, which are crucial in the context of a Bayesian formulation.
3. Here, we only place ourselves in the case of VTI profile. No azimuthal anisotropy is considered. Furthermore, it is well known that lateral heterogeneities (e.g., Moho topography) may also produce apparent radial anisotropy. This case has not been considered here.

4.7 Conclusion

Global tomographic imaging is an inverse problem where different types of observables (e.g., surface waves, body waves) are used to constrain different types of parameters (e.g., P and S wave velocity, anisotropy, density). Different measurements have different sensitivities, and all parameters are not equally well resolved. Some parameters present strong trade-offs. Furthermore, the elastic properties to be constrained are scale dependent. These issues have led seismologists to simplify the inverse problem and to invert separately different observables, different frequency bands, and with different spatial and physical parameterizations. Thus, different classes of models with different resolving power have been published, which are sometimes difficult to reconcile.

Recent theoretical developments as well as increased availability of computational power have allowed the emergence of full waveform inversion, where the full wave-field (and its derivatives with respect to the model) is exactly computed with purely numerical methods. Inverting the full wave-field in the time domain enables us to combine body and surface waves in the same inversion scheme. The recent local, regional, and global applications of full waveform tomography reveal indeed an improved resolution. For example, the amplitudes of the imaged anomalies increase, and more small-scale features are constrained. However, a remaining challenge is computational cost. One way to keep reasonable the time of computations is to limit the frequency content of waveforms and only invert the long periods. As a result, the resolving power of full waveform tomography is mainly determined by the minimum period or minimum wavelength.

Elastic properties are scale dependent, and hence, the small-scale heterogeneities that are not resolved in tomographic models are mapped into large-scale structure. This mapping is nonlinear, which makes large-scale structure difficult to interpret. In this study, we have focused on vertical transversely isotropic (VTI)

models, where small-scale heterogeneities in S wave velocity are mapped into large-scale radial anisotropy.

A method has been proposed to "downscale" smooth tomographic models. Given the laws of homogenization which link a medium to its effective long-wave equivalent, we are able to explore the space of possible small-scale models that are equivalent to a given tomographic profile. We cast this inverse problem in a Bayesian formulation, which enables us to probabilistically quantify the trade-off between heterogeneity (in our case layering) and anisotropy.

We have applied this method to a tomographic profile of the North American craton, and added short-period information from receiver functions, to help locate the depth of discontinuities. This allows us to investigate the nature and history of the craton. But more importantly, we have shown that a large part of anisotropy present in the tomographic model may be due to unmapped discontinuities. Therefore, we conclude that one should not interpret radial anisotropy in tomographic models only in terms of geodynamics, e.g., mantle flow. The inferred radial anisotropy contains valuable information about the earth, but one has to keep in mind that this is only apparent anisotropy. It may only be interpreted when associated with higher frequency information, or with additional information from other disciplines (geology, mineral physics, or geodynamics).

Here, we have assumed that the observed radial anisotropy is either due to layering or intrinsic radial anisotropy. We recognize that this a strong approximation. Indeed, both lateral heterogeneities and intrinsic azimuthal anisotropy may also contribute to the observed radial anisotropy. Although these effects are not considered here, they are the subject of current work. For example, we expect observations of azimuthal anisotropy to provide important additional constraints: (1) These observations cannot be explained by horizontal layering; (2) they may sometimes be due to LPO with a tilted axis of symmetry, in which case part of the observed radial anisotropy may still be intrinsic.

Acknowledgments We thank Andreas Fichtner and Lapo Boschi for providing constructive reviews. This project was supported by the French-Berkeley Fund (FBF). Thomas Bodin wishes to acknowledge support from the Miller Institute for Basic Research at the University of California, Berkeley. Y. Capdeville was partly supported by the French ANR blanche "mm" (grant ANR-10-BLAN-613 MEME). Computations were performed on the CCIPL computer "Erdre." Computational resources were also provided by the Institutional Computing program of LANL. We are grateful to Cheng Cheng and Scott French for providing figures. We also thank Scott French for providing model SEMum2. We are also very grateful to Heather Ford and Karen Fischer for providing teleseismic P waveforms recorded at station YKW3.

References

Abt D, Fischer K, French S, Ford H, Yuan H, Romanowicz B (2010) North american lithospheric discontinuity structure imaged by PS and SP receiver functions. J Geophys Res 115(10.1029)

Aki K (1982) Scattering and attenuation. Bull Seismol Soc Am 72(6B):S319–S330

Aki K, Christoffersson A, Husebye ES (1977) Determination of the three-dimensional seismic structure of the lithosphere. J Geophys Res 82(2):277–296

Allaire G (1992) Homogenization and two-scale convergence. SIAM J Math Anal 23:1482–1518
Anderson DL (1961) Elastic wave propagation in layered anisotropic media. J Geophys Res 66(9):2953–2963
Anderson DL, Dziewonski AM (1982) Upper mantle anisotropy: evidence from free oscillations. Geophys J Int 69(2):383–404
Auriault J-L, Sanchez-Palencia E (1977) Étude du comportement macroscopique d'un milieu poreux saturé déformable. J Mécanique 16(4):575–603
Auriault J-L, Borne L, Chambon R (1985) Dynamics of porous saturated media, checking of the generalized law of Darcy. J Acoust Soc Am 77:1641
Babuška V, Cara M (1991) Seismic anisotropy in the earth, vol 10. Springer, Berlin
Backus G, Gilbert F (1968) The resolving power of gross earth data. Geophys J Roy Astron Soc 16(2):169–205
Backus GE (1962) Long-wave elastic anisotropy produced by horizontal layering. J Geophys Res 67(11):4427–4440
Bayes T (1763) An essay towards solving a problem in the doctrine of chances, C. Davis, Printer to the Royal Society of London
Bensen G, Ritzwoller M, Yang Y (2009) A 3-d shear velocity model of the crust and uppermost mantle beneath the united states from ambient seismic noise. Geophys J Int 177(3):1177–1196
Bensoussan A, Lions J-L, Papanicolaou G (1978) Asymptotic analysis of periodic structures, North Holland. AMS Chelsea Publishing, Providence
Bercovici D, Karato S-I (2003) Whole-mantle convection and the transition-zone water filter. Nature 425(6953):39–44
Bijwaard H, Spakman W, Engdahl E (1998) Closing the gap between regional and global travel time tomography. J Geophys Res 103(B12):30055
Bodin T, Maupin V (2008) Resolution potential of surface wave phase velocity measurements at small arrays. Geophys J Int 172(2):698–706
Bodin T, Salmon M, Kennett B, Sambridge M (2012a) Probabilistic surface reconstruction from multiple data-sets-an example for the Australian moho. J Geophys Res 117(B10)
Bodin T, Sambridge M, Tkalčić H, Arroucau P, Gallagher K, Rawlinson N (2012b) Transdimensional inversion of receiver functions and surface wave dispersion. J Geophys Res 117:B02301
Bodin T, Yuan H, Romanowicz B (2014) Inversion of receiver functions without deconvolution application to the Indian craton. Geophys J Int 196(2):1025–1033
Bonnet M, Constantinescu A (2005) Inverse problems in elasticity. Inverse Prob 21(2):R1
Bostock M (1998) Mantle stratigraphy and evolution of the slave province. J Geophys Res Solid Earth (1978–2012) 103(B9):21183–21200
Bostock M (2014) Teleseismic body-wave scattering and receiver-side structure. Treatise Geophys 1:219–246
Box G, Tiao G (1973) Bayesian Inference in Statistical Inference. Addison-Wesley, Boston
Bozdağ E, Trampert J (2008) On crustal corrections in surface wave tomography. Geophys J Int 172(3):1066–1082
Burdick L, Langston C (1977) Modeling crustal structure through the use of converted phases in teleseismic body-wave forms. Bull Seismol Soc Am 67(3):677–691
Burdick S, Li C, Martynov V, Cox T, Eakins J, Mulder T, Astiz L, Vernon FL, Pavlis GL, van der Hilst RD (2008) Upper mantle heterogeneity beneath North America from travel time tomography with global and USArray transportable array data. Seismol Res Lett 79(3):384–392
Capdeville Y, Marigo J-J (2007) Second order homogenization of the elastic wave equation for non-periodic layered media. Geophys J Int 170(2):823–838
Capdeville Y, Cance P (2015) Residual homogenization for elastic wave propagation in complex media.Geophys J Int 200(2):984−997
Capdeville Y, To A, Romanowicz B (2003) Coupling spectral elements and modes in a spherical earth: an extension to the sandwich case. Geophys J Int 154(1):44–57

Capdeville Y, Gung Y, Romanowicz B (2005) Towards global earth tomography using the spectral element method: a technique based on source stacking. Geophys J Int 162(2):541–554
Capdeville Y, Guillot L, Marigo J-J (2010a) 1-D non-periodic homogenization for the seismic wave equation. Geophys J Int 181(2):897–910
Capdeville Y, Guillot L, Marigo J-J (2010b) 2-D non-periodic homogenization to upscale elastic media for p–sv waves. Geophys J Int 182(2):903–922
Capdeville Y, Zhao M, Cupillard P (2015) Fast fourier homogenization for elastic wave propagation in complex media. Wave Motion
Capdeville Y, Stutzmann É, Wang N, Montagner J-P (2013) Residual homogenization for seismic forward and inverse problems in layered media. Geophys J Int 194(1):470–487
Carcione JM (1994) The wave equation in generalized coordinates. Geophysics 59(12):1911–1919
Chapman C (2004) Fundamentals of seismic wave propagation. Cambridge University Press, Cambridge
Chen C-W, Rondenay S, Weeraratne DS, Snyder DB (2007) New constraints on the upper mantle structure of the slave craton from rayleigh wave inversion. Geophys Res Lett 34(10):L10301
Chen C-W, Rondenay S, Evans RL, Snyder DB (2009) Geophysical detection of relict metasomatism from an Archean (3.5 ga) subduction zone. Science 326(5956):1089–1091
Cheng C, Allen RM, Porritt RW, Ballmer MD (2015) Seismic constraints on a double-layered asymmetric whole-mantle plume beneath Hawaii. Hawaiian volcanoes: from source to surface. pp 19–34
Choblet G, Husson L, Bodin T (2014) Probabilistic surface reconstruction of coastal sea level rise during the twentieth century 2014. J Geophys Res (Solid Earth) 119:9206–9236. doi:10.1002/2014JB011639
Colli L, Fichtner A, Bunge H-P (2013) Full waveform tomography of the upper mantle in the south Atlantic region: imaging a westward fluxing shallow asthenosphere? Tectonophysics 604:26–40
Cook FA, Velden AJ, Hall KW, Roberts BJ (1999) Frozen subduction in Canada's Northwest Territories: lithoprobe deep lithospheric reflection profiling of the western Canadian shield. Tectonics 18(1):1–24
Crampin S (1981) A review of wave motion in anisotropic and cracked elastic-media. Wave Motion 3(4):343–391
Crampin S, Booth DC (1985) Shear-wave polarizations near the north Anatolian fault–II. Interpretation in terms of crack-induced anisotropy. Geophys J Int 83(1):75–92
Dahlen F, Hung S-H, Nolet G (2000) Fréchet kernels for finite-frequency traveltimes–I. Theory. Geophys J Int 141(1):157–174
Debayle E, Kennett B (2000) Anisotropy in the Australasian upper mantle from love and rayleigh waveform inversion. Earth Planet Sci Lett 184(1):339–351
Debayle E, Ricard Y (2012) A global shear velocity model of the upper mantle from fundamental and higher rayleigh mode measurements. J Geophys Res Solid Earth (1978–2012) 117(B10)
Dettmer J, Dosso S, Holland C (2010) Trans-dimensional geoacoustic inversion. J Acoust Soc Am 128:3393
Domingos P (1999) The role of Occam's razor in knowledge discovery. Data Min Knowl Disc 3(4):409–425
Dziewonski A, Anderson D (1981) Preliminary reference earth model. Phys Earth Planet Inter 25(4):297–356
Ekström G, Dziewonski AM (1998) The unique anisotropy of the pacific upper mantle. Nature 394(6689):168–172
Endrun B, Meier T, Lebedev S, Bohnhoff M, Stavrakakis G, Harjes H-P (2008) S velocity structure and radial anisotropy in the aegean region from surface wave dispersion. Geophys J Int 174(2):593–616
Ferreira A, Woodhouse J, Visser K, Trampert J (2010) On the robustness of global radially anisotropic surface wave tomography. J Geophys Res Solid Earth (1978–2012) 115(B4)
Fichtner A, Igel H (2008) Efficient numerical surface wave propagation through the optimization of discrete crustal models a technique based on non-linear dispersion curve matching (dcm). Geophys J Int 173(2):519–533

Fichtner A, Trampert J (2011) Resolution analysis in full waveform inversion. Geophys J Int 187(3):1604–1624
Fichtner A, Bunge H-P, Igel H (2006) The adjoint method in seismology: I. Theory. Phys Earth Planet Inter 157(1):86–104
Fichtner A, Kennett BL, Igel H, Bunge H-P (2009) Full seismic waveform tomography for upper-mantle structure in the australasian region using adjoint methods. Geophys J Int 179(3):1703–1725
Fichtner A, Kennett BL, Igel H, Bunge H-P (2010) Full waveform tomography for radially anisotropic structure: new insights into present and past states of the australasian upper mantle. Earth Planet Sci Lett 290(3):270–280
Fichtner A, Kennett BL, Trampert J (2013a) Separating intrinsic and apparent anisotropy. Phys Earth Planet Inter 219:11–20
Fichtner A, Trampert J, Cupillard P, Saygin E, Taymaz T, Capdeville Y, Villaseñor A (2013b) Multiscale full waveform inversion. Geophys J Int 194(1):534–556. doi:10.1093/gji/ggt118
Ford HA, Fischer KM, Abt DL, Rychert CA, Elkins-Tanton LT (2010) The lithosphere–asthenosphere boundary and cratonic lithospheric layering beneath australia from Sp wave imaging. Earth Planet Sci Lett 300(3):299–310
French S, Lekic V, Romanowicz B (2013) Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere. Science 342(6155):227–230
Friederich W, Huang Z-X (1996) Evidence for upper mantle anisotropy beneath southern Germany from love and rayleigh wave dispersion. Geophys Res Lett 23(10):1135–1138
Fukao Y, Obayashi M (2013) Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity. J Geophys Res Solid Earth 118(11):5920–5938
Fukao Y, Widiyantoro S, Obayashi M (2001) Stagnant slabs in the upper and lower mantle transition region. Rev Geophys 39(3):291–323
Geyer C, Møller J (1994) Simulation procedures and likelihood inference for spatial point processes. Scand J Stat 21(4):359–373
Gilbert F, Dziewonski A (1975) An application of normal mode theory to the retrieval of structural parameters and source mechanisms for seismic spectra. Philos Trans R Soc London A 278:187–269
Green P (1995) Reversible jump MCMC computation and Bayesian model selection. Biometrika 82:711–732
Green P (2003) Trans-dimensional Markov chain Monte Carlo. Highly Struct Stoch Syst 27:179–198
Guillot L, Capdeville Y, Marigo J-J (2010) 2-D non-periodic homogenization of the elastic wave equation: Sh case. Geophys J Int 182(3):1438–1454
Gung Y, Panning M, Romanowicz B (2003) Global anisotropy and the thickness of continents. Nature 422(6933):707–711
Hashin Z, Shtrikman S (1963) A variational approach to the elastic behavior of multiphase materials. J Mech Phys Solids 11:127–140
Hastings W (1970) Monte Carlo simulation methods using Markov chains and their applications. Biometrika 57:97–109
Helbig K, Thomsen L (2005) 75-plus years of anisotropy in exploration and reservoir seismics: a historical review of concepts and methods Geophysics 70(6):9ND–23ND
Hess H (1964) Seismic anisotropy of the uppermost mantle under oceans. Nature 203:629–631
Hopper E, Ford HA, Fischer KM, Lekic V, Fouch MJ (2014) The lithosphere–asthenosphere boundary and the tectonic and magmatic history of the northwestern united states. Earth Planet Sci Lett 402:69–81
Houser C, Masters G, Shearer P, Laske G (2008) Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms. Geophys J Int 174(1):195–212
Iaffaldano G, Bodin T, Sambridge M (2012) Reconstructing plate-motion changes in the presence of finite-rotations noise. Nat Commun 3:1048
Iaffaldano G, Bodin T, Sambridge M (2013) Slow-downs and speed-ups of India–Eurasia convergence since: data-noise, uncertainties and dynamic implications. Earth Planet Sci Lett 367:146–156

Kárason H, Van Der Hilst RD (2000) Constraints on mantle convection from seismic tomography. Hist Dyn Glob Plate Mot 277–288
Kawakatsu H, Kumar P, Takei Y, Shinohara M, Kanazawa T, Araki E, Suyehiro K (2009) Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. Science 324(5926):499–502
Kelly K, Ward R, Treitel S, Alford R (1976) Synthetic seismograms: a finite-difference approach. Geophysics 41(1):2–27
Kennett B (1995) Approximations for surface-wave propagation in laterally varying media. Geophys J Int 122(2):470–478
Kennett B, Nolet G (1990) The interaction of the S-wavefield with upper mantle heterogeneity. Geophys J Int 101(3):751–762
Khan A, Boschi L, Connolly J (2011) Mapping the earth's thermochemical and anisotropic structure using global surface wave data. J Geophys Res Solid Earth (1978–2012) 116(B1)
Kind R, Yuan X, Kumar P (2012) Seismic receiver functions and the lithosphere–asthenosphere boundary. Tectonophysics 536:25–43
Komatitsch D, Tromp J (1999) Introduction to the spectral element method for three-dimensional seismic wave propagation. Geophys J Int 139(3):806–822
Komatitsch D, Tromp J (2002) Spectral-element simulations of global seismic wave propagation—I. Validation. Geophys J Int 149(2):390–412
Komatitsch D, Vilotte J-P (1998) The spectral element method: an efficient tool to simulate the seismic response of 2d and 3d geological structures. Bull Seismol Soc Am 88(2):368–392
Komatitsch D, Vinnik LP, Chevrot S (2010) Shdiff-svdiff splitting in an isotropic earth. J Geophys Res 115(B7):B07312
Kustowski B, Ekström G, Dziewoński A (2008) Anisotropic shear-wave velocity structure of the earth's mantle: a global model. J Geophys Res Solid Earth (1978–2012) 113(B6)
Langston C (1979) Structure under Mount Rainier, Washington, inferred from teleseismic body waves. J Geophys Res 84(B9):4749–4762
Lebedev S, Van Der Hilst RD (2008) Global upper-mantle tomography with the automated multimode inversion of surface and s-wave forms. Geophys J Int 173(2):505–518
Lebedev S, Boonen J, Trampert J (2009) Seismic structure of Precambrian lithosphere: new constraints from broad-band surface-wave dispersion. Lithos 109(1):96–111
Lekić V, Romanowicz B (2011) Inferring upper-mantle structure by full waveform tomography with the spectral element method. Geophys J Int 185(2):799–831
Lekić V, Panning M, Romanowicz B (2010) A simple method for improving crustal corrections in waveform tomography. Geophys J Int 182(1):265–278
Lekic V, Cottaar S, Dziewonski A, Romanowicz B (2012) Cluster analysis of global lower mantle tomography: a new class of structure and implications for chemical heterogeneity. Earth Planet Sci Lett 357:68–77
Levshin A, Ratnikova L (1984) Apparent anisotropy in inhomogeneous media. Geophys J Int 76(1):65–69
Li C, van der Hilst RD, Engdahl ER, Burdick S (2008) A new global model for P wave speed variations in earth's mantle. Geochem Geophys Geosyst 9(5)
Li X-D, Romanowicz B (1995) Comparison of global waveform inversions with and without considering cross-branch modal coupling. Geophys J Int 121(3):695–709
Li X-D, Romanowicz B (1996) Global mantle shear velocity model developed using nonlinear asymptotic coupling theory. J Geophys Res Solid Earth (1978–2012) 101(B10):22245–22272
Li X-D, Tanimoto T (1993) Waveforms of long-period body waves in a slightly aspherical earth model. Geophys J Int 112(1):92–102
Liu Q, Gu Y (2012) Seismic imaging: from classical to adjoint tomography. Tectonophysics 566:31–66
Love A (1927) The mathematical theory of elasticity. CRC Press, Boca Raton
Mainprice D (2007) Seismic anisotropy of the deep earth from a mineral and rock 1022 physics perspective. In: Schubert G (ed) Treatise in geophysics, vol 2, pp 437–492

Malinverno A (2002) Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. Geophys J Int 151(3):675–688
Maupin V (2001) A multiple-scattering scheme for modelling surface wave propagation in isotropic and anisotropic three-dimensional structures. Geophys J Int 146(2):332–348
Maupin V, Park J (2014) Theory and observations—wave propagation in anisotropic media. Seismol Struct Earth Treatise Geophys 1:289–321
Mégnin C, Romanowicz B (2000) The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms. Geophys J Int 143(3):709–728
Metropolis N et al (1953) Equations of state calculations by fast computational machine. J Chem Phys 21(6):1087–1091
Montagner J (2014) Upper mantle structure: global isotropic and anisotropic tomography. Treatise Geophys 1:559–590
Montagner J-P (1994) Can seismology tell us anything about convection in the mantle? Rev Geophys 32(2):115–137
Montagner J-P, Anderson DL (1989) Petrological constraints on seismic anisotropy. Phys Earth Planet Inter 54(1):82–105
Montagner J-P, Guillot L (2002) Seismic anisotropy and global geodynamics. Rev Mineral Geochem 51(1):353–385
Montagner J-P, Jobert N (1988) Vectorial tomography—II. Application to the Indian ocean. Geophys J Int 94(2):309–344
Montagner J-P, Nataf H-C (1988) Vectorial tomography—I. Theory. Geophys J 94(2):295–307
Murat F, Tartar L (1985) Calcul des variations et homogénéisation. In: Homogenization methods: theory and applications in physics (Bréau-sans-Nappe, 1983), vol 57 of Collect. Dir. Études Rech. Élec. France, Eyrolles, Paris, pp 319–369
Muyzert E, Paulssen H, Snieder R (1999) A seismic cross-section through the east European continent. Geophys J Int 136(3):695–704
Nachman AI (1988) Reconstructions from boundary measurements. Ann Math 128(3):531–576
Nakamura G, Uhlmann G (1994) Global uniqueness for an inverse boundary problem arising in elasticity. Inventiones mathematicae 118(1):457–474
Nolet G (1990) Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs. J Geophys Res Solid Earth (1978–2012) 95(B6):8499–8512
Obrebski M, Allen RM, Pollitz F, Hung S-H (2011) Lithosphere–asthenosphere interaction beneath the western united states from the joint inversion of body-wave traveltimes and surface-wave phase velocities. Geophys J Int 185(2):1003–1021
Panning M, Romanowicz B (2006) A three-dimensional radially anisotropic model of shear velocity in the whole mantle. Geophys J Int 167(1):361–379
Panning M, Lekić V, Romanowicz B (2010) Importance of crustal corrections in the development of a new global model of radial anisotropy. J Geophys Res Solid Earth (1978–2012) 115(B12)
Park M, Odom RI (1999) The effect of stochastic rough interfaces on coupled-mode elastic waves. Geophys J Int 136(1):123–143
Pilia S, Rawlinson N, Cayley RA, Bodin T, Musgrave R, Reading AM, Direen NG, Young MK (2015) Evidence of micro-continent entrainement during crustal accretion. Scientific Reports, NPG, 5:8218. doi:10.1038/srep08218
Postma G (1955) Wave propagation in a stratified medium. Geophysics 20(4):780–806
Priolo E, Carcione JM, Seriani G (1994) Numerical simulation of interface waves by high-order spectral modeling techniques. J Acoust Soc Am 95(2):681–693
Rawlinson N, Sambridge M (2003) Seismic traveltime tomography of the crust and lithosphere. Adv Geophys 46:81–198
Rawlinson N, Kennett BLN, Salmon M, Glen RA (2015) Origin of Lateral Heterogeneities in the Upper Mantle Beneath South-east Australia from Seismic Tomography, this volume
Ray A, Key K (2012) Bayesian inversion of marine CSEM data with a trans-dimensional self parametrizing algorithm. Geophys J Int 191:1135–1151

Raykova RB, Nikolova SB (2003) Anisotropy in the earth's crust and uppermost mantle in southeastern Europe obtained from rayleigh and love surface waves. J Appl Geophys 54(3):247–256

Ricard Y, Durand S, Montagner J-P, Chambat F (2014) Is there seismic attenuation in the mantle? Earth Planet Sci Lett 388:257–264

Richards PG, Menke W (1983) The apparent attenuation of a scattering medium. Bull Seismol Soc Am 73(4):1005–1021

Rickers F, Fichtner A, Trampert J (2013) The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion. Earth Planet Sci Lett 367:39–51

Ritsema J, van Heijst HJ, Woodhouse JH (1999) Complex shear wave velocity structure imaged beneath Africa and iceland. Science 286(5446):1925–1928

Ritsema J, Deuss A, Van Heijst H, Woodhouse J (2011) S40rts: a degree-40 shear-velocity model for the mantle from new rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. Geophys J Int 184(3):1223–1236

Romanowicz B (1987) Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the earth's normal modes. Geophys J Roy Astron Soc 90(1):75–100

Romanowicz B (2002) Inversion of surface waves: a review. Int Geophys 81A:149–173

Romanowicz B (2003) Global mantle tomography: progress status in the past 10 years. Ann Rev Earth Planet Sci 31(1):303–328

Romanowicz BA, Panning MP, Gung Y, Capdeville Y (2008) On the computation of long period seismograms in a 3-d earth using normal mode based approximations. Geophys J Int 175(2):520–536

Sambridge M, Bodin T, Gallagher K, Tkalčić H (2013) Transdimensional inference in the geosciences. Philos Trans R Soc A Math Phys Eng Sci 371(1984)

Sanchez-Palencia E (1980) Non homogeneous media and vibration theory, no. 127 in Lecture Notes in Physics. Springer, Berlin

Schaeffer A, Lebedev S (2013) Global shear speed structure of the upper mantle and transition zone. Geophys J Int 194(1):417–449

Schaeffer AJ, Lebedev S (2015) Global Heterogeneity of the Lithosphere and Underlying Mantle: A Seismological Appraisal Based on Multimode Surface-Wave Dispersion Analysis, Shear-Velocity Tomography, and Tectonic Regionalization, this volume

Schmerr N (2015) Imaging Mantle Heterogeneity with Upper Mantle Seismic Discontinuities, this volume

Seriani G, Priolo E (1994) Spectral element method for acoustic wave simulation in heterogeneous media. Finite Elem Anal Des 16(3):337–348

Shearer P (1991) Imaging global body wave phases by stacking long-period seismograms. J Geophys Res 96(B12):20353–20364

Sheng P (1990) Effective-medium theory of sedimentary rocks. Phys Rev B 41(7):4507

Sigloch K, Mihalynuk MG (2013) Intra-oceanic subduction shaped the assembly of cordilleran North America. Nature 496(7443):50–56

Sisson S (2005) Transdimensional Markov Chains: a decade of progress and future perspectives. J Am Stat Assoc 100(471):1077–1090

Sivia D (1996) Data analysis: a Bayesian tutorial. Oxford University Press, USA

Tape C, Liu Q, Maggi A, Tromp J (2010) Seismic tomography of the southern california crust based on spectral-element and adjoint methods. Geophys J Int 180(1):433–462

Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. Society for Industrial Mathematics

Tessmer E, Kosloff D, Behle A (1992) Elastic wave propagation simulation in the presence of surface topography. Geophys J Int 108(2):621–632

Thomson WT (1950) Transmission of elastic waves through a stratified solid medium. J Appl Phys 21:89

Tkalčić H, Young M, Bodin T, Ngo S, Sambridge M (2013) The shuffling rotation of the earth/'s inner core revealed by earthquake doublets. Nat Geosci 6:497–502

Trampert J, Woodhouse JH (2003) Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s. Geophys J Int 154(1):154–165

Tromp J, Tape C, Liu Q (2005) Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophys J Int 160(1):195–216

Van Der Lee S, Frederiksen A (2005) Surface wave tomography applied to the north american upper mantle. Geophys Monogr Ser 157:67–80

Van der Lee S, Nolet G (1997) Upper mantle s velocity structure of North America. J Geophys Res 102(B10):22815–22822

Vinnik L (1977) Detection of waves converted from P to SV in the mantle. Phys Earth Planet Inter 15(1):39–45

Vinnik LP, Farra V, Romanowicz B (1989) Azimuthal anisotropy in the earth from observations of SKS at Geoscope and NARS broadband stations. Bull Seismol Soc Am 79(5):1542–1558

Virieux J (1986) P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method. Geophysics 51(4):889–901

Virieux J, Operto S (2009) An overview of full-waveform inversion in exploration geophysics. Geophysics 74(6):WCC1–WCC26

Wang N, Montagner J-P, Fichtner A, Capdeville Y (2013) Intrinsic versus extrinsic seismic anisotropy: the radial anisotropy in reference earth models. Geophys Res Lett 40(16):4284–4288

Woodhouse JH, Dziewonski AM (1984) Mapping the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms. J Geophys Res Solid Earth (1978–2012) 89(B7):5953–5986

Young MK, Rawlinson N, Bodin T (2013) Transdimensional inversion of ambient seismic noise for 3d shear velocity structure of the tasmanian crust. Geophysics 78(3):WB49–WB62

Yuan H, Romanowicz B (2010) Lithospheric layering in the North American craton. Nature 466(7310):1063–1068

Yuan X, Kind R, Li X, Wang R (2006) The s receiver functions: synthetics and data example. Geophys J Int 165(2):555–564

Zhu H, Tromp J (2013) Mapping tectonic deformation in the crust and upper mantle beneath Europe and the north atlantic ocean. Science 341(6148):871–875

Zhu H, Bozda E, Peter D, Tromp J (2012) Structure of the European upper mantle revealed by adjoint tomography. Nat Geosci 5(7):493–498