Geophys. J. Int. (2008) **173,** 519–533 doi: 10.1111/j.1365-246X.2008.03746.x

Efficient numerical surface wave propagation through the optimization of discrete crustal models—a technique based on non-linear dispersion curve matching (DCM)

Andreas Fichtner and Heiner Igel

Department of Earth and Environmental Sciences, Geophysics Section, Ludwig-Maximilian University, Munich, Germany.
E-mail: andreas.fichtner@geophysik.uni-muenchen.de

Accepted 2008 January 28. Received 2008 January 25; in original form 2007 November 8

SUMMARY

We present a method for reducing the computational costs of numerical surface wave modelling. It is based on the smoothing of thin near-surface layers and discontinuities. Optimal smooth models are found via a constrained non-linear matching of dispersion curves in the period range of interest. The major advantages of our method are that it is independent of the numerical techniques employed and that it does not require modifications of pre-existing codes or meshes. It is, moreover, applicable in cases where the layer thickness is of the order of one wavelength and automatically yields estimates of the appropriateness of the smoothed model. Even though our analysis is based on 1-D media, we demonstrate with a numerical example that the dispersion curve matching can yield satisfactory results when it is applied regionally to laterally heterogeneous models. Also in that case it can lead to considerable reductions of the computational requirements.

Key words: Numerical solutions; Tomography; Surface waves and free oscillations; Seismic anisotropy; Computational seimology; Wave propagation.

1 INTRODUCTION

The simulation of seismic wave propagation through realistic earth models is becoming increasingly feasible and necessary as both data quality and computational resources improve. All numerical methods that find approximate solutions of the wave equation rely on the replacement of a continuously defined earth model $\oplus$ by a discrete model $\oplus_{\mathrm{d}}$. Due to this discretization, small-scale structures from $\oplus$, such as thin layers and discontinuities, can often not be represented accurately in $\oplus_{\mathrm{d}}$ unless the grid spacing is reduced. This can lead to large numerical errors, especially in the surface wave trains. Our aim therefore, is to deduce discrete versions of models with thin layers and discontinuities that give accurate numerical solutions without reducing the grid spacing.

For the discrete representation of discontinuities from $\oplus$, several method-specific solutions already exist. All of them lead to an increase of the computational costs or the algorithmic complexity or both. Here we will mention only two of many numerical techniques used for elastic wave propagation: (1) the spectral-element method (SEM) and (2) the finite-difference method (FDM). The SEM (e.g. Priolo *et al.* 1994; Faccioli *et al.* 1997; Seriani 1998; Chaljub & Valette 2004) is based on the weak form of the equations of motion and a decomposition of the computational domain into disjoint elements. Inside the elements, the dynamic fields are approximated by high-order polynomials that are necessarily smooth. This implies that material discontinuities need to coincide with element boundaries for the solution to be correct. If, however, a discontinuity is located inside an element—due to limited computational resources, for example—then the rapidly varying fields can not be represented accurately. The results are of such an inaccurate model discretization are unacceptably large numerical errors. A reduction of the grid spacing—either globally or locally through non-conforming grids (e.g. Chaljub *et al.* 2003)—can, in principle, eliminate these errors. However, the increasing computational costs can render this option impossible. Work by Kelly *et al.* (1976) indicated early on that the FDM also requires material discontinuities to be treated explicitly. The FDM (e.g. Virieux 1986; Igel *et al.* 2002) directly discretizes the strong form of the equations of motion by replacing derivatives with difference quotients. Since the strong form of the wave equation is valid only for continuously varying media, boundary conditions should be imposed explicitly along discontinuities to achieve optimal accuracy (e.g. Moczo *et al.* 2002).

Thin layers present more general difficulties than isolated discontinuities, because their width may be smaller than the smallest computationally feasible grid spacing. Backus (1962) and Schoenberg & Muir (1989) demonstrated that a stack of thin layers can be replaced by an equivalent slowly varying earth model. Their analysis is based on the assumption that the widths of the individual layers are much smaller

than the dominant wavelength and that the elastic deformations can be treated as static. The equivalent smooth earth model was shown to be transversely isotropic even when the original stack of layers is isotropic. The long-wavelength equivalence of smooth earth models can clearly be used for the benefit of numerical efficiency because modest variations of the model parameters require neither reductions of the grid spacing nor additional algorithmic complexities. A concept similar to the one advocated by Backus (1962) and Schoenberg & Muir (1989) is the homogenization of the equations of motion with rapidly varying coefficients (e.g. Capdeville & Marigo 2007; Stuart & Pavliotis 2007). This technique explicitly yields a smoothed model by solving the so-called cell problem, composed of a partial differential equation and a set of subsidiary conditions. Again, the solutions for the slowly varying model are close to the correct solution when the length scale of the variations in $\oplus$ is much smaller than the dominant wavelength.

Here we present an alternative approach for constructing smoothed and long-wavelength equivalent earth models that can be discretized through direct sampling. This helps to overcome difficulties with the discrete representation of discontinuities and thin layers as long as a pre-defined maximum frequency is not exceeded. The construction of the smoothed earth models is based on the matching of surface wave dispersion curves corresponding to the original model $\oplus$ and a smoothed model $\oplus_s$. This means that we search in a pool of smooth models for the one that best reproduces the dispersion curves of Love and Rayleigh waves in $\oplus$. The considered frequency band is the one for which a given grid can, in principle, yield accurate numerical solutions. Body waves are not explicitly taken into account because, in the considered period range, they are comparatively insensitive to crustal structure.

This paper is organized as follows. First, we illustrate the effects of improperly implemented crustal discontinuities on the quality of numerically computed surface waves. A slowly varying and long-wavelength equivalent model is then presented in Section 3. Subsequently, we elaborate on the methodology, covering issues such as the construction of the smooth models, their dynamic stability and uniqueness. Finally, the advantages and limitations of our method are addressed in the discussion.

The broader objective of our study is to contribute to the advancement of waveform tomography on continental and global scales (e.g. Capdeville *et al.* 2005; Sieminski *et al.* 2007), for which accurately modelled surface waves are indispensable.

2 ILLUSTRATION OF THE PROBLEM

We start by illustrating the effects of an improperly implemented crustal model on the accuracy of numerically computed waveforms. The example is intended to serve as a motivation for the subsequent developments and as a means for introducing concepts that we shall use throughout this study. To solve the elastic wave equation numerically, we employ a spectral-element method, henceforth referred to as ses3d. It operates in a 3-D spherical section with absorbing boundaries. The numerical grid is equidistant in the radial and the angular directions. Even though we focus our attention on the spectral-element method, our principal results also apply to other techniques, which solve the equations of motion by discretizing the computational domain.

First, we demonstrate the accuracy of the ses3d solutions by comparing them with semi-analytic solutions. The earth model is spherically symmetric with a homogeneous, single-layered crust. It is identical to PREM (Dziewonski & Anderson 1981) below the depth of 30 km. The width of the elements is chosen such that the crust is represented by exactly one layer of elements (Fig. 1). The black dots in Fig. 1 represent the irregularly spaced Gauss–Lobatto–Legendre points which serve as gridpoints in the SEM.

The source is located at 80 km depth and the cut-off period of the seismograms is 15 s. We compute the reference seismograms using the program package GEMINI, developed and described in detail by Friederich & Dalkolmo (1995). It allows us to compute accurate waveforms for both body and surface waves, keeping the computational costs low. The results for a receiver located at an epicentral distance of 21.3° are shown in the top row of Fig. 2. Qualitatively, the GEMINI synthetics (dotted curve) and the ses3d synthetics (solid curve) are in excellent agreement for all three components.

It is important for our purposes to assess the differences between the semi-analytic and the numerical solutions quantitatively also. For this we closely follow the suggestions by Kristeková *et al.* (2006) and consider envelope and phase misfits. The analysis is based on the

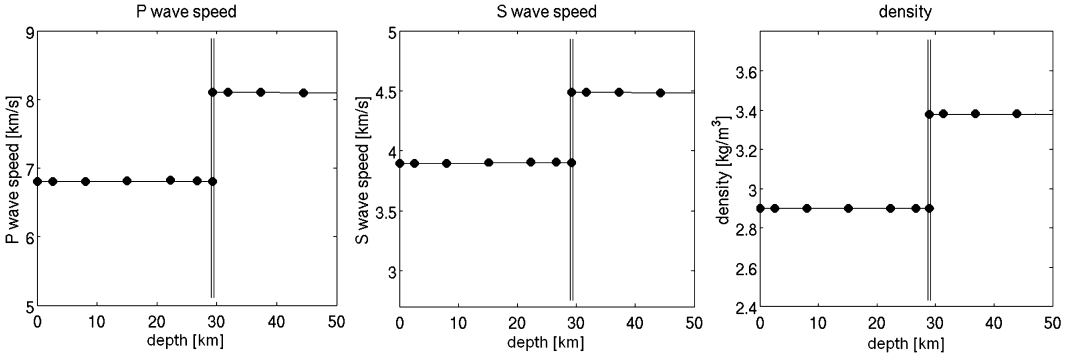

Figure 1. Details of the vertical discretization of the homogeneous crust overlying PREM. Dots symbolize the SEM gridpoints (Gauss–Lobatto–Legendre points) obtained by sampling the continuously defined model plotted as a solid line. One layer of elements, each comprising 7^3 points, is used to represent the crust. The vertical double line marks the boundary between the two upper layers of elements.

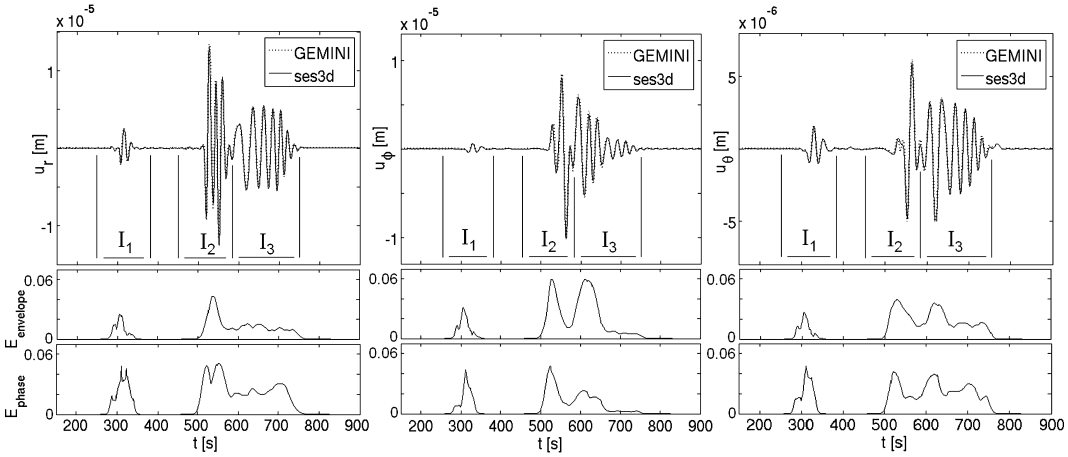

Figure 2. Top row: Comparison of synthetic seismograms computed with GEMINI (dotted curve) and ses3d (solid curve) for the homogeneous crust model (Fig. 1) and a station at an epicentral distance of 21.3°. The time intervals for which the envelope and phase misfits have been computed separately are indicated by I_1 (*P*-wave arrivals), I_2 (*S*-wave arrivals) and I_3 (surface wave train). Bottom rows: The envelope and phase misfits, as defined in the text. The envelope misfit is mostly below 5 per cent and the phase misfit rarely exceeds the value of 0.05 rad.

time–frequency representation $\hat{f}(\omega, t)$ of a time-series $f(t)$ defined as

$$\hat{f}(\omega, t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) h_\sigma(\tau - t)\, \mathrm{e}^{-i\omega\tau}\, \mathrm{d}\tau, \tag{1}$$

where h_σ denotes a positive window function. We have chosen h_σ to be the Gaussian

$$h_\sigma(t) = (\pi\sigma^2)^{-1/4} \mathrm{e}^{-t^2/(2\sigma^2)}, \tag{2}$$

so that $\hat{f}(\omega, t)$ is the Gabor transform of f. The Gabor transform has the advantageous property of maximizing the time–frequency resolution. A useful value for the parameter σ is the dominant wavelength of the signal, which is in our case 15 s. The necessity of choosing a time window can be circumvented by using the continuous wavelet transform instead of the Gabor transform (Kristeková *et al.* 2006). For our purposes, however, the Gabor transform is fully sufficient. A detailed treatment of time–frequency analysis can, for example, be found in Strang & Nguyen (1996). Denoting by u_i^G and u_i^s the *i*-components of the GEMINI and the ses3d synthetics, respectively, their instantaneous phase difference $\phi_i^s(\omega, t) - \phi_i^G(\omega, t)$ can be expressed in terms of their time–frequency representations:

$$\phi_i^s(\omega, t) - \phi_i^G(\omega, t) = -i\, \mathrm{Ln} \left[\frac{\hat{u}_i^s(\omega, t)\, |\hat{u}_i^G(\omega, t)|}{\hat{u}_i^G(\omega, t)\, |\hat{u}_i^s(\omega, t)|} \right]. \tag{3}$$

The symbol *Ln* denotes the principal value logarithm, that is, the branch where the phase ranges between $-\pi$ and π. A phase misfit E_p that quantifies phase differences as a function of time can then be defined as

$$E_p(t)\big|_{t\in I_k} := \frac{\int_{-\infty}^{\infty} |\hat{u}_i^s(\omega, t)| [\phi_i^s(\omega, t) - \phi_i^G(\omega, t)]\,\big|_{t\in I_k}\, \mathrm{d}\omega}{\max_{t\in I_k} \int_{-\infty}^{\infty} |\hat{u}_i^s(\omega, t)|\, \mathrm{d}\omega}. \tag{4}$$

The phase misfit is evaluated separately in three intervals comprising the P arrivals (I_1), the S arrivals (I_2) and the surface wave train (I_3) (see Fig. 2). Large phase differences in high-amplitude waveforms generate large phase misfits whereas equally large phase differences in low-amplitude waveforms produce small values of E_p. The emphasis is, therefore, on the large amplitudes. In analogy to the definition of the phase misfit, one can define the envelope misfit E_e:

$$E_e(t) := \frac{\int_{-\infty}^{\infty} \big\| \hat{u}_i^s(\omega, t)| - |\hat{u}_i^G(\omega, t) \big\|\big|_{t\in I_k}\, \mathrm{d}\omega}{\max_{t\in I_k} \int_{-\infty}^{\infty} |\hat{u}_i^s(\omega, t)|\, \mathrm{d}\omega}. \tag{5}$$

Both, E_p and E_e are meaningful only when the two solutions are at least similar. Extreme differences between u_i^G and u_i^s can not be quantified properly with E_p and E_s. Note that the quantification of the differences between the GEMINI and the ses3d solutions is not meant to be a quantification of one of the method's accuracy. It is instead intended to serve as a diagnostic tool and a detector of inaccuracies that would sometimes be difficult to quantify by a purely visual comparison of seismograms. The phase and envelope misfits for our initial example with the single-layered crust are plotted in the bottom rows of Fig. 2. Outside the intervals I_1, I_2, I_3 we do not compute misfits, which does not mean that they are not there. Mostly the envelope misfits do not exceed 5 per cent and the phase misfits are mostly below 0.05 rad. This translates to a time-shift of roughly 0.12 s, which is slightly more than the time step of 0.1 s used for the SEM simulation.

The usefulness of the time–frequency domain misfits E_p and E_e as diagnostic tools becomes apparent when we change the earth model so that the crust now comprises two layers, as shown in Fig. 3. Two discontinuities are now located inside the upper layer of elements. It would, in this particular example, be possible to adapt the numerical grid to the new model by reducing the vertical grid spacing non-uniformly so that one layer of elements coincides with each of the crustal layers. Though this would in principle be feasible, the case of multiple layers with varying thicknesses would clearly result in strongly increased computational costs.

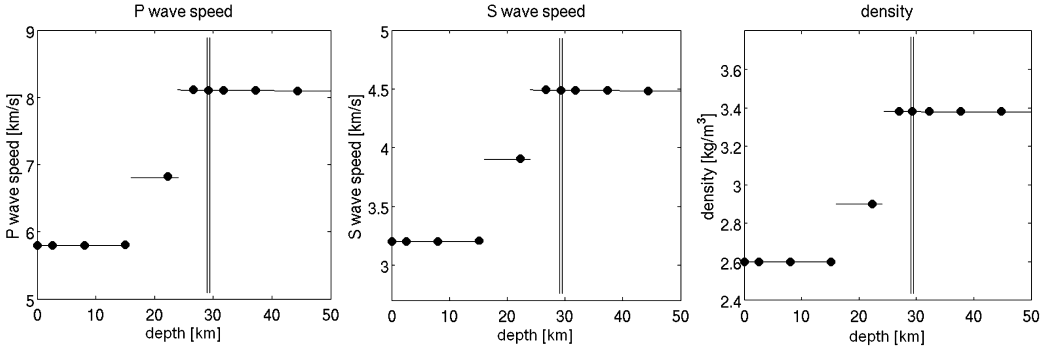

Figure 3. Details of the vertical discretization of the crust comprising two layers. Dots symbolize the SEM gridpoints (Gauss–Lobatto–Legendre points). Two discontinuities are now located inside the upper layer of elements. There, the SEM displacement and stress fields are represented by smooth polynomials even though the exact solution varies rapidly due to reflection and refraction phenomena. The double vertical line marks the boundary between the two upper layers of elements.

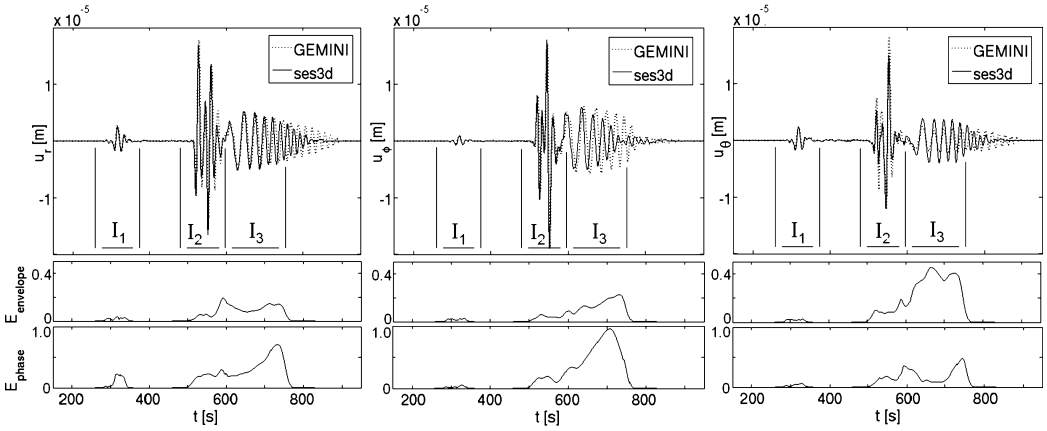

Figure 4. Top panel: Comparison of synthetic seismograms computed with GEMINI (dotted line) and ses3d (solid line) for the two-layer crust model (Fig. 3) and a station at an epicentral distance of 21.3°. While the inadequate implementation of the crust has little effect on the body waves, the surface waves are severely affected. The ses3d surface waves notably differ from the semi-analytic GEMINI solution. Bottom panel: Amplitude and phase misfits corresponding to the above seismograms. Misfits for $t > 750$ s are not computed because the differences between u_i^G and u_i^s are too large to be properly expressible through E_p and E_e.

Therefore, we deliberately do not adopt this option. Instead, we use this example to illustrate and quantify the numerical errors arising from such an improper model implementation. Since the rapidly varying wavefield in the vicinity of the crustal discontinuities can not be represented by the smooth polynomials inside the elements we can expect larger discrepancies between the GEMINI and the ses3d solutions. That they indeed occur, can be seen in Fig. 4. Though the body waves remain largely unaffected, the phase and envelope misfits in the surface wave train become unacceptably high. They reach values of 40 per cent for the envelope misfit and 1.0 rad for the phase misfit. The latter corresponds to time-shifts of almost 3 s. Again, we did not compute the misfits outside the intervals I_1, I_2, I_3. In fact, for $t > 750$ s, the discrepancies between the two solutions are so large that E_p and E_e are not meaningful anymore.

3 CAN WE FIND SMOOTH CRUSTAL MODELS?

We want to tackle the problem illustrated above by replacing the upper part of the original earth model by an equivalent smooth model. For this, we exploit the fact that surface waves in a limited frequency range do not uniquely determine an earth model. That is, there is a non-empty set of earth models, all of which produce surface wave solutions that differ from each other by less than a subjectively chosen small value. From this set of models we can, therefore, choose one, which is numerically advantageous, that is, smooth.

To find such a smooth model, one should focus on those aspects of the surface wave train that are most important in the real data analysis and that facilitate the model construction. Therefore, we propose to construct smooth earth models by matching the surface wave dispersion curves of the original model with the dispersion curves of models where the upper part is defined in terms of low-order polynomials. In the general case, the smoothed models will be transversely isotropic (see Backus 1962). The minimization of the dispersion curve misfit is done non-linearly by simulated annealing (SA) (Kirkpatrick *et al.* 1983). To improve the uniqueness of the smooth models we impose that they be close to the original earth model. Before delving into the methodological details, we first justify our approach by its results shown in Figs 5–8.

A slowly varying crustal model that reproduces almost exactly the surface wave dispersion curves from the original model, is displayed in Fig. 5. That the model is mildly anisotropic inside the upper layer of elements can be seen from the differences between the wave speeds of *SH*

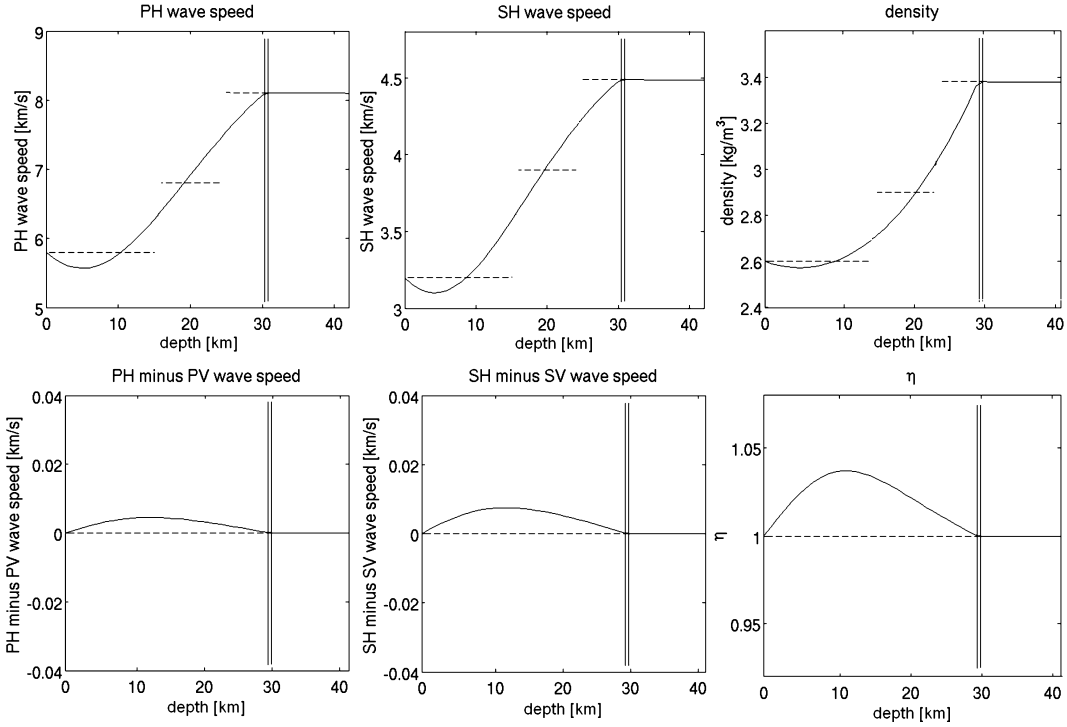

Figure 5. Upper part of the original earth model (dashed line) with two crustal layers (Fig. 3) and the smoothed earth model (solid line). The smooth model is slightly anisotropic with radial symmetry axis. Therefore, the wave speeds of *SH* and *SV* waves do not coincide, just as the wave speeds of *PH* and *PV* waves. Also, the parameter η (see e.g. Takeuchi & Saito 1972; Dziewonski & Anderson 1981) is different from 1. The smooth model is given in terms of 4th order polynomials. For Love waves with periods between 16 and 67 s, the average phase velocity misfit of the fundamental mode is 2.5×10^{-4} km s^{-1}. The corresponding misfit for Rayleigh waves is 3.9×10^{-4} km s^{-1}. Hence, the surface wave dispersion curves for the original model and the smoothed model are practically indistinguishable between 16 and 67 s. (See also Fig. 6.)

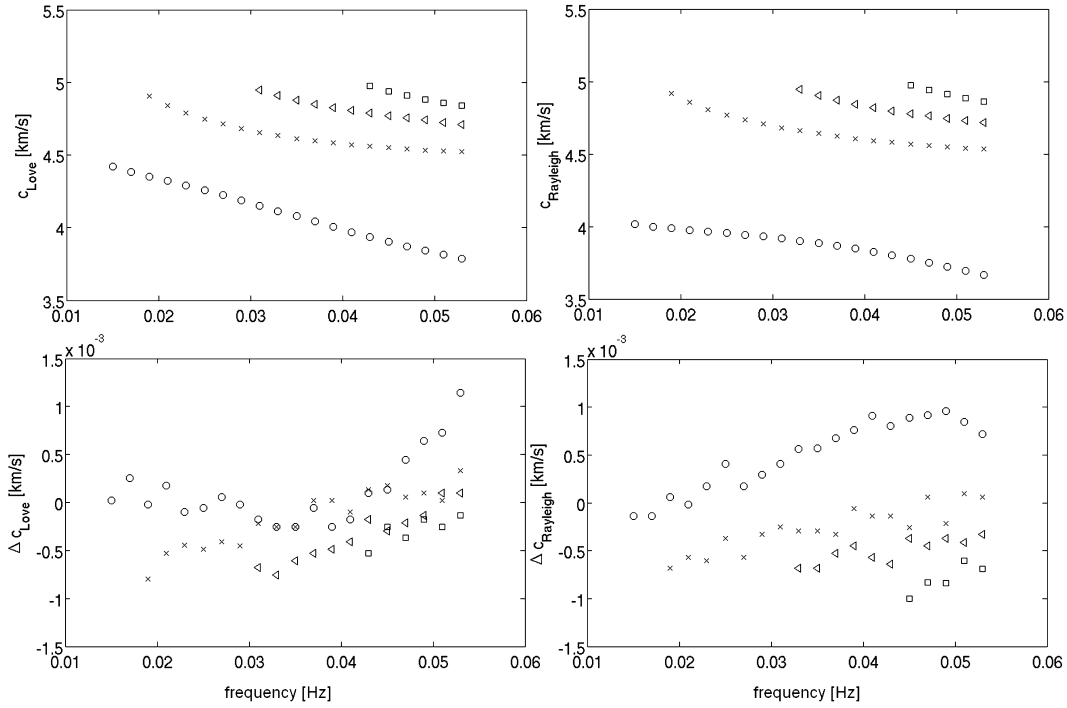

Figure 6. Top panel: Dispersion curves for Love and Rayleigh waves for the original earth model. Circles denote the respective fundamental modes and crosses (×), triangles (◁) and squares (□) represent the higher modes. Bottom panel: Corresponding dispersion curve errors (smoothed minus original) that are due to the smoothing of the crust are shown.

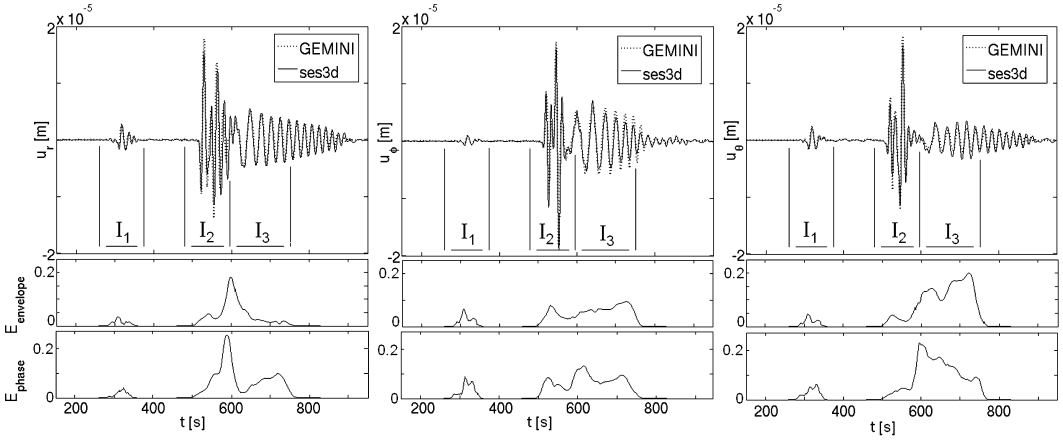

Figure 7. Top panel: Comparison of synthetic seismograms computed with GEMINI (dotted line) and ses3d (solid line) for the same station considered in Fig. 4. The GEMINI seismograms were computed for the earth model with the layered crust (dashed lines in Fig. 5) and the ses3d seismograms were computed for the corresponding smoothed model (solid lines in Fig. 5). The surface wave part is now well modelled. Bottom panel: Corresponding envelope and phase misfits. Both misfits are significantly smaller than in the case of the inaccurately implemented crust shown in Fig. 4. We did not compute misfits for $t > 750$ s to be consistent with Fig. 4.

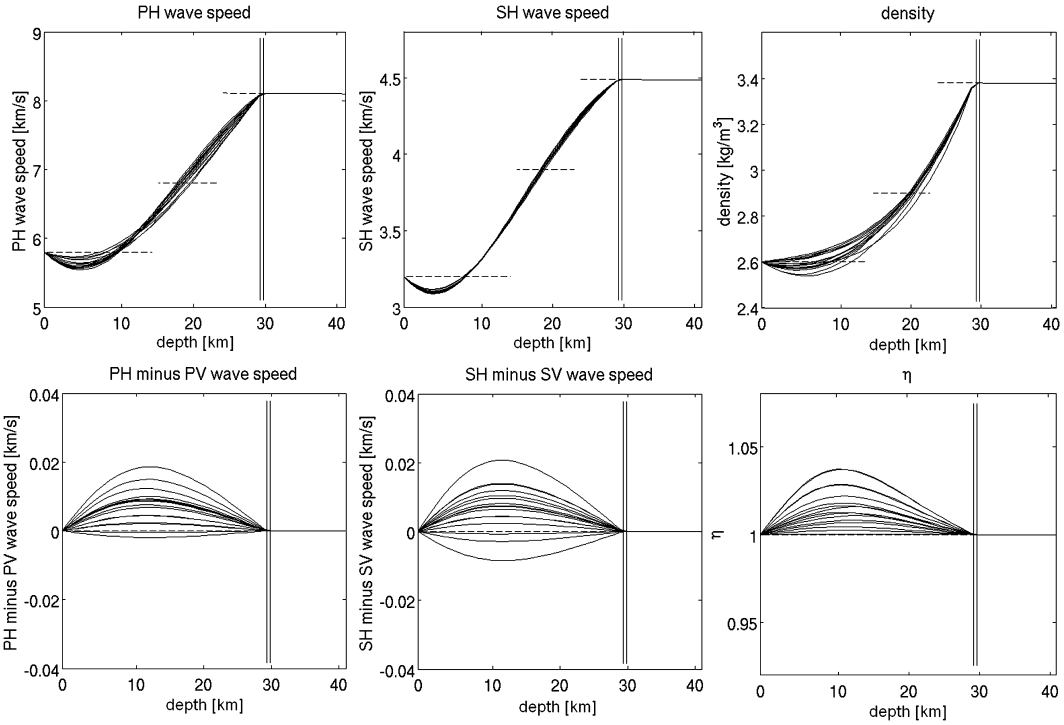

Figure 8. Display of 20 smoothed models found by jointly minimizing dispersion curves and model differences, as described in the text. The dashed lines represent the respective original models. All models produce average dispersion curve misfits below 7×10^{-4} km s^{-1} for both fundamental Love and Rayleigh waves with periods between 17 and 67 s.

and *SV* waves, and the wave speeds of *PH* and *PV* waves. Here, *PH* refers to a horizontally travelling *P* wave and *PV* to a vertically travelling *P* wave. Moreover, the parameter η (see e.g. Takeuchi & Saito 1972; Dziewonski & Anderson 1981, or the section on the methodology) is different from 1. For the fundamental modes of Love and Rayleigh waves, the mean difference in the phase velocities for the two models is less than 4×10^{-4} km s^{-1} in the period range 17–67 s. Details of the dispersion curves and the associated misfits are shown in Fig. 6. For both Love and Rayleigh waves the absolute phase velocity differences rarely exceed 1×10^{-3} km s^{-1} and are mostly below 0.5×10^{-3} km s^{-1}. The lower limit of the considered period range (17 s) has been chosen such that it coincides with the shortest periods in the synthetic seismograms. Periods above the upper limit (67 s) do not contribute significantly in the construction of the smooth models. They have therefore, been disregarded.

That the new model does indeed remove most of the discrepancies between the semi-analytic and the SEM solution can be seen in Fig. 7. Both the envelope and the phase misfits are now close to the ones observed in the case of the correctly implemented single-layered crustal model. This demonstrates that the small dispersion curve misfit (Fig. 6) is indeed a reliable indicator of significantly smaller numerical errors.

Since the construction of the slowly varying crustal model is based on an SA algorithm, the result is inherently non-unique. A whole family of models can be expected to produce essentially the same surface wave dispersion characteristics, at least in a limited period range. This non-uniqueness can be reduced significantly by requiring that the smooth model be close to the original one. This additional constraint is also in the interest of accurately modelled amplitudes. In the context of our method, uniqueness is, in principle, not required. It is, however, intuitively desirable. Moreover, the degree to which a parameter of the smooth model is non-unique—despite additional constraints—sheds some light onto the physics of wave propagation and the necessity of anisotropy. In Fig. 8 we display 20 smoothed models, all of which produce average dispersion curve misfits below 7×10^{-4} km s^{-1} for both fundamental Love and Rayleigh waves with periods between 17 and 67 s. The smoothed *SH* wave speed profile is stronger constrained than the *PH* wave speed because the latter is allowed to vary more widely while keeping the dispersion curve misfit acceptably small. The density profile and the anisotropy (bottom row in Figs 5 and 8) are the least constrained.

4 METHODOLOGY

The modification of discrete earth models for the benefit of numerical accuracy can be interpreted as the optimization of a discretized differential operator. Let us denote by $L_{ij}\mathbf{u}_j = \mathbf{f}_i$, a discretized version of the wave equation $L\mathbf{u} = \mathbf{f}$, where L (L_{ij}) is the (discrete) wave equation operator, $\mathbf{u}$ ($\mathbf{u}_j$) the (discrete) displacement field and $\mathbf{f}$ ($\mathbf{f}_i$) the (discrete) external force density. We are interested in finding L_{ij} such that $\mathbf{u}_j$ is as close as possible to the exact solution $\mathbf{u}$.

Generally, the discrete operator L_{ij} depends on both the numerical method and the discrete version $\oplus_{\mathrm{d}}$ of the continuously defined earth model $\oplus$. Since we wish to leave the numerical method unchanged, we are left with the option of modifying $\oplus_{\mathrm{d}}$ to improve the numerical solution.

Ideally, one would directly apply the numerical method—in our case the SEM—to a set of smoothed discrete test models. One could then choose the discrete model that best satisfies the strong criterion that the wavefields $\mathbf{u}_j(\oplus_d)$ and $\mathbf{u}(\oplus)$ be close to each other. In practice, however, exact solutions $\mathbf{u}(\oplus)$ are often unavailable and solving the discrete problem sufficiently many times is computationally too expensive. Therefore, we propose to replace this strong and direct criterion by a weaker and indirect criterion: We require that $\oplus$ and a smooth and continuously defined earth model $\oplus_{\mathrm{s}}$ produce essentially identical surface wave dispersion curves in the period range of interest. Then $\oplus_{\mathrm{s}}$ is sampled at the numerical gridpoints to give the discrete model $\oplus_{\mathrm{d}}$. Since $\oplus_{\mathrm{s}}$ is by construction slowly varying, it is represented accurately by $\oplus_{\mathrm{d}}$. Consequently, the surface wave dispersion of the numerical solution $\mathbf{u}_j(\oplus_d)$ is close to the surface wave dispersion of the exact solution $\mathbf{u}(\oplus)$—assuming, of course, that the errors introduced by the numerical method itself are comparatively small. This approach is justified by the observation that merely sampling $\oplus$ to obtain $\oplus_{\mathrm{d}}$ leads to unacceptably large numerical errors, mainly in the surface wave trains. That dispersion curve matching indeed reduces the numerical errors significantly has already been demonstrated by comparison with semi-analytic solutions.

We will give the details of our method in the following paragraphs, starting with spherically symmetric earth models. In the interest of completeness and clear symbolisms, we briefly review the equations of motion for Love and Rayleigh waves in an 1-D and transversely isotropic earth model. We then elaborate on the SA algorithm used for the minimization of the differences between two sets of dispersion curves and on the random generation of earth models. Based on the algorithms developed for spherically symmetric models, we finally consider earth models with lateral heterogeneities.

4.1 The equations of motion for Love and Rayleigh waves in layered media

For the first part of the analysis we assume that the earth model $\oplus$ is spherically symmetric. The computation of surface wave dispersion curves in such a 1-D medium is based on the equations of motion for Love and Rayleigh waves. Since we are primarily interested in periods below 50 s the analysis will be in Cartesian coordinates. As demonstrated by Backus (1962) the smoothing of thin layers leads to an apparent anisotropy with vertical symmetry axis, even when the individual layers are isotropic. We therefore, consider a transversely isotropic medium with an elastic tensor given by

$$C = \begin{pmatrix} c_{zzzz} & c_{zzyy} & c_{zzxx} & c_{zzyx} & c_{zzzx} & c_{zzzy} \\ c_{yyzz} & c_{yyyy} & c_{yyxx} & c_{yyyx} & c_{yyzx} & c_{yyzy} \\ c_{xxzz} & c_{xxyy} & c_{xxxx} & c_{xxyx} & c_{xxzx} & c_{xxzy} \\ c_{yxzz} & c_{yxyy} & c_{yxxx} & c_{yxyx} & c_{yxzx} & c_{yxzy} \\ c_{zxzz} & c_{zxyy} & c_{zxxx} & c_{zxyx} & c_{zxzx} & c_{zxzy} \\ c_{zyzz} & c_{zyyy} & c_{zyxx} & c_{zyyx} & c_{zyzx} & c_{zyzy} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda+2\mu & \lambda+c & \lambda+c & 0 & 0 & 0 \\ \lambda+c & \lambda+2\mu+a & \lambda+a & 0 & 0 & 0 \\ \lambda+c & \lambda+a & \lambda+2\mu+a & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu+b & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu+b \end{pmatrix}. \tag{6}$$

The elastic coefficients are chosen such that an anisotropic medium appears as a perturbation of an isotropic medium in which $a = b = c = 0$. The elastic parameters λ, μ, a, b and c are related to the classical ones A, C, F, L and N—introduced by Love (1892)—by $A = \lambda + 2\mu + a$, $C = \lambda + 2\mu$, $F = \lambda + c$, $L = \mu + b$ and $N = \mu$. Moreover, we have $\rho v_{ph}^2 = \lambda + 2\mu + a$, $\rho v_{pv}^2 = \lambda + 2\mu$, $\rho v_{\mathrm{sh}}^2 = \mu$ and $\rho v_{\mathrm{sv}}^2 = \mu + b$. The dimensionless parameter η (see e.g. Takeuchi & Saito 1972; Dziewonski & Anderson 1981) is related to λ, a and c via $\eta = (\lambda + c)/(\lambda + a)$. Under the assumption of a vertically stratified medium, *SH* and *P–SV* waves can be treated separately, leading to independent Love and Rayleigh waves, respectively. Following Takeuchi & Saito (1972) and Aki & Richards (2002) we choose the ansatz

$$u_x = 0, \quad u_y = l_1(k, \omega, z)\, e^{\mathbf{i}(kx-\omega t)}, \quad u_z = 0, \tag{7}$$

for the plane Love wave displacement in the spatio–temporal frequency domain. The symbols k and ω denote the horizontal wave number and the temporal frequency, respectively. Upon inserting eqs (6) and (7) into the non-dissipative elastic wave equation, we obtain a first-order system for the displacement–stress vector $\mathbf{l} = (l_1, l_2)^T$:

$$\frac{\mathrm{d}}{\mathrm{d}z} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = \begin{pmatrix} 0 & (\mu+b)^{-1} \\ k^2\mu - \rho\omega^2 & 0 \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix}. \tag{8}$$

The solutions of eq. (8) are subject to the boundary condition $l_2|_{z=0} = 0$ (free surface) and the radiation condition $\lim_{z\to\infty} = 0$. In analogy to eq. (7) the plane Rayleigh wave displacement can be written as

$$u_x = r_1(k, \omega, z)\, e^{\mathbf{i}(kx-\omega t)}, \quad u_y = 0, \quad u_z = \mathbf{i} r_2(k, \omega, z)\, e^{\mathbf{i}(kx-\omega t)}. \tag{9}$$

The corresponding displacement–stress vector $\mathbf{r} = (r_1, r_2, r_3, r_4)^T$ is then determined by the first-order system

$$\frac{\mathrm{d}}{\mathrm{d}z} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = \begin{pmatrix} 0 & k & (\mu+b)^{-1} & 0 \\ -k(\lambda+c)(\lambda+2\mu)^{-1} & 0 & 0 & (\lambda+2\mu)^{-1} \\ k^2\xi\,(\lambda+2\mu)^{-1} - \rho\omega^2 & 0 & 0 & k(\lambda+c)(\lambda+2\mu)^{-1} \\ 0 & -\rho\omega^2 & -k & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}, \tag{10}$$

with $\xi := 4\mu(\lambda+\mu) + a(\lambda+2\mu) - c(2\lambda+c)$. The free surface conditions for the system (eq. 10) are $r_3|_{z=0} = r_4|_{z=0} = 0$, and the radiation conditions are $\lim_{z\to\infty} r_1 = \lim_{z\to\infty} r_2 = 0$. Two independent solutions for $\mathbf{r}$ can be obtained by starting the integration at a sufficiently large depth where $r_1 = r_2 = 0$ and by setting $(r_3, r_4) = (1, 0)$ and $(r_3, r_4) = (0, 1)$, respectively. labeling these two solutions by $\mathbf{r}^{(1)}$ and $\mathbf{r}^{(2)}$, the free surface condition transforms to

$$r_3^{(1)} r_4^{(2)} - r_3^{(2)} r_4^{(1)}\big|_{z=0} = 0. \tag{11}$$

The stress vectors $r_3^{(1)}, r_4^{(1)}, r_3^{(2)}$ and $r_4^{(2)}$ attain very large values, especially as the frequency increases. This leads to inaccurate results because very large numbers need to be subtracted to find the zero in eq. (11). Therefore, the differential equation for the displacement–stress vector (eq. 10) is usually rewritten in terms of the six minors

$$\begin{aligned} R_1 &:= r_1^{(1)} r_2^{(2)} - r_2^{(1)} r_1^{(2)}, \quad R_2 := r_1^{(1)} r_3^{(2)} - r_3^{(1)} r_1^{(2)}, \\ R_3 &:= r_1^{(1)} r_4^{(2)} - r_4^{(1)} r_1^{(2)}, \quad R_4 := r_2^{(1)} r_3^{(2)} - r_3^{(1)} r_2^{(2)}, \\ R_5 &:= r_2^{(1)} r_4^{(2)} - r_4^{(1)} r_2^{(2)}, \quad R_6 := r_3^{(1)} r_4^{(2)} - r_4^{(1)} r_3^{(2)}. \end{aligned} \tag{12}$$

They can then be shown to satisfy the equations

$$\frac{\mathrm{d}}{\mathrm{d}z} R_1 = -(\mu+b)^{-1} R_4 + (\lambda+2\mu)^{-1} R_3, \tag{13a}$$

$$R_2 = -R_5, \tag{13b}$$

$$\frac{\mathrm{d}}{\mathrm{d}z} R_3 = 2k R_5 + (\mu+b)^{-1} R_6 - \rho\omega^2 R_1, \tag{13c}$$

$$\frac{\mathrm{d}}{\mathrm{d}z} R_4 = 2k(\lambda+c)(\lambda+2\mu)^{-1} R_5 - [k^2\xi\,(\lambda+2\mu)^{-1} - \rho\omega^2]\, R_1 - (\lambda+2\mu)^{-1} R_6, \tag{13d}$$

$$\frac{\mathrm{d}}{\mathrm{d}z} R_5 = -k(\lambda+c)(\lambda+2\mu)^{-1} R_3 - k R_4, \tag{13e}$$

$$\frac{\mathrm{d}}{\mathrm{d}z} R_6 = [k^2 \xi\, (\lambda + 2\mu)^{-1} - \rho\omega^2] R_3 + \rho\omega^2 R_4. \tag{13f}$$

The free surface condition is now reduced to $R_6|_{z=0} = 0$. To find the dispersion relations $k = k(\omega)$, we fix ω and then determine the values of k for which solutions of the systems eqs (8) and (13), with their respective subsidiary conditions, exist.

4.2 Non-linear dispersion curve matching by Simulated Annealing

Let us denote by $c_L^i(\omega, \oplus)$ and $c_R^i(\omega, \oplus)$ the frequency-dependent phase velocities of the ith mode of the Love and Rayleigh waves in the earth model $\oplus = (\rho, \mu, \lambda, a, b, c)$. Our aim is to find a continuously defined smooth model $\oplus_s = (\rho_s, \mu_s, \lambda_s, a_s, b_s, c_s)$ such that

$$\begin{aligned} E = &\sum_{i=0}^{n} \alpha_i \int_{\omega=\omega_1}^{\omega_2} \left| c_L^i(\omega, \oplus) - c_L^i(\omega, \oplus_s) \right| \mathrm{d}\omega \\ &+ \sum_{i=0}^{n} \alpha_i \int_{\omega=\omega_1}^{\omega_2} \left| c_R^i(\omega, \oplus) - c_R^i(\omega, \oplus_s) \right| \mathrm{d}\omega + \beta \,||\oplus - \oplus_s||_M \end{aligned} \tag{14}$$

is minimized. The misfit measure E includes the dispersion curve misfits for Rayleigh and Love waves in a pre-defined frequency interval $[\omega_1\ \omega_2]$ and the relative L_1 distance between the two models $\oplus$ and $\oplus_s$. We decided to use the L_1 norm because it produced the best results in the numerical experiments. Typically, n is 3 or 4. Including higher modes does not improve the results very much because they are mostly sensitive to deeper structures. The numbers α_i and β are weighting factors.

Including the term $||\oplus - \oplus_s||_M$, and therefore, forcing the smooth model to be close to the original model, serves a variety of purposes. First, it reduces the non-uniqueness of the problem. Uniqueness is in principle not required but still intuitively desirable. Keeping the differences between $\oplus$ and $\oplus_s$ as small as possible is also important in the context of smoothing earth models with lateral heterogeneities, as will be discussed in Section 4.5. In the presence of sources inside the smoothed region, one would in principle have to modify the moment tensor. This complication can be avoided when $||\oplus - \oplus_s||_M$ is small.

Minimizing E with respect to $\oplus_s$ is a non-linear problem that can be solved by a variety of standard techniques. Since the calculation of surface wave dispersion curves in a 1-D medium is computationally inexpensive, we decided to employ a SA algorithm, introduced by Kirkpatrick *et al.* (1983). In addition to being easy to implement, SA also offers the possibility to impose constraints on the solution simply by means of random model generation. Implementing constraints in a gradient-based minimization would be more involved. This is the case especially when the constraints are in the form of inequalities, such as the ones deduced in Section 4.4 (stability constraints). Moreover, there is no guarantee that gradient methods lead to the global minimum.

4.3 Random model generation

The SA algorithm requires random test models that are compared with the current best model. To generate test models that serve their purpose, we use two different parametrizations: In the lower part of the model where smoothing is unnecessary, we retain the original model parametrization. In the upper part of the model, however, we employ low-order piecewise Lagrange polynomials, defined on the elements used for the SEM simulations. The collocation points are the Gauss–Lobatto points. Therefore, the models vary more rapidly near the edges than the centres of the elements. This corresponds well to the parametrization of the displacement and stress fields.

Depending on the numerical method used, different sets of collocation points may work equally well or better. For example, in the case of a classical finite-difference method with equally spaced gridpoints, it would be natural to choose equally spaced collocation points for a polynomial model representation.

Random models are then generated through the perturbation of the polynomial coefficients. To ensure that the test models are physically reasonable, we impose the following conditions: (1) The smoothed model and the original model are identical at the surface. This restriction is motivated by the fact that the amplitudes of seismic waves are very sensitive to the structure directly beneath the receiver. This worked well in our example, but there are scenarios where this restriction can be impractical. One of them is a thin, very low-velocity layer at the top of the model (e.g. sediments). There, one might rather set the velocity at the surface equal to the velocity of a deeper layer with higher velocities. (2) There are no jumps at the element boundaries. Hence, there are no additional discontinuities that may generate undesirable reflections. (3) The smoothed mass density distribution is strictly positive, that is, $\rho_s > 0$. (4) As we will demonstrate in the next paragraph, we also need to impose $\mu_s \geq 0$, $\lambda_s \geq 0$, $|a_s| \geq \mu_s/2$, $|b_s| \geq \mu_s/2$ and $|c_s| \geq \mu_s/2$ to ensure that the smooth earth model be stable.

4.4 Stability considerations

The parameters of the slowly varying earth model cannot be chosen arbitrarily. They have to satisfy a set of conditions to be physically admissible. In addition to the obvious requirement that the smoothed density distribution ρ_s be strictly positive ($\rho_s > 0$), the parameters λ_s, μ_s, a_s, b_s and c_s need to conform to the stability criterion. Stability means that the EARTH can either oscillate about its equilibrium configuration (zero displacement) or that it can be slightly perturbed away from the equilibrium configuration, without changing the potential energy. A system that is physically unstable will not be numerically stable unless the numerical error somehow compensates the physical

instability. In mathematical terms, stability is equivalent to the requirement that the 6×6 matrix $\boldsymbol{C}_s$ (eq. 6) be positive semi-definite. The block structure of $\boldsymbol{C}_s$ then implies that the inequalities $\mu_s \geq 0$ and $\mu_s + b_s \geq 0$ must hold and that the 3×3 submatrix

$$\tilde{\boldsymbol{C}}_s = \begin{pmatrix} \lambda_s + 2\mu_s & \lambda_s + c_s & \lambda_s + c_s \\ \lambda_s + c_s & \lambda_s + 2\mu_s + a_s & \lambda_s + a_s \\ \lambda_s + c_s & \lambda_s + a_s & \lambda_s + 2\mu_s + a_s \end{pmatrix} \tag{15}$$

be positive semi-definite. Instead of computing the eigenvalues of $\tilde{\boldsymbol{C}}_s$ we invoke Sylvester's criterion, which states that a symmetric matrix is positive semi-definite if and only if all its principal minors are non-negative. After some algebraic manipulations, we find that the complete set of inequalities that need to be satisfied is

$$\mu_s \geq 0, \quad \mu_s + b_s \geq 0, \quad \lambda_s + 2\mu_s \geq 0, \quad \lambda_s + \mu_s + a_s \geq 0, \quad (\lambda_s + 2\mu_s)(\lambda_s + \mu_s + a_s) \geq (\lambda_s + c_s)^2. \tag{16}$$

Note that in the case of isotropy, that is for $a_s = b_s = c_s = 0$, the stability conditions reduce to $\mu_s \geq 0$ and $-1 \leq \nu_s \leq 1/2$, where $\nu_s = \lambda_s/(2\lambda_s + 2\mu_s)$ is Poisson's ratio in the smoothed medium. It is convenient and sufficient for our purposes to impose more restrictive conditions that automatically imply that the inequalities (16) hold:

$$\mu_s \geq 0, \quad \lambda_s \geq 0, \quad |a_s| \leq \mu_s/2, \quad |b_s| \leq \mu_s/2, \quad |c_s| \leq \mu_s/2. \tag{17}$$

The requirement that $\lambda_s \geq 0$ is equivalent to saying that Poisson's ratio be positive. This is a physically reasonable assumption for earth materials, which are unlikely to be auxetic.

4.5 Heterogeneous earth models

The very purpose of numerical wave propagation is the computation of wavefields in heterogeneous earth models for which analytic solutions are mostly unavailable. So far, however, our analysis has been restricted to purely spherically symmetric earth models $\oplus^{1-\mathrm{D}}$ because dispersion curves cannot be computed exactly in the presence of lateral heterogeneities. A direct application of the DCM to an arbitrary model $\oplus = \oplus^{1-\mathrm{D}} + \delta\oplus$, is therefore, not possible. For the construction of smoothed models in the 3-D, case one has to make the distinction between lateral heterogeneities that are large or small compared with the wavelength of the elastic waves. We will consider the two cases separately:

4.5.1 Long-wavelength heterogeneities

In the case of long-wavelength lateral heterogeneities—present for example in the crustal model Crust2.0 (Bassin *et al.* 2000)—the DCM can be applied regionally. This approach is illustrated in a 3-D example where the structural model is divided into an oceanic and a continental part. In the continental crust, the *P*- and *S*-wave speeds vary laterally by ± 2.5 and ± 5.0 per cent, respectively. Significant lateral variations occur over length scales of 1° (see Fig. 9). Since the model is 3-D, we compute a reference solution using our spectral-element code ses3d. The thickness of the oceanic crust, and therefore, the width of the elements is 9.1 km. Also the boundary between the upper and lower continental crusts is honoured. With this implementation, we can simulate surface waves with periods that are as short as 5 s—much shorter than the periods typically used in continental-scale surface wave analysis. The details of this setup are summarized in Table 1.

We construct the 3-D long-wavelength equivalent model as follows. First, we compute a smooth crustal model for the oceanic part and a smooth crustal model for the average continental part. Those, together with the original model, are shown in Fig. 9. The width of the new elements is chosen to be 28 km, which is equal to the depth of the lower continental crust. Hence, the wider elements do not honour the discontinuities at the bottom of the oceanic crust and between the upper and lower continental crusts. In the case of the oceanic crust, there is nothing more to do because it does not vary laterally. The continental crust is subdivided into regions where the *S* wave speed varies between -5 and -4, -4 and -3 per cent, ..., $+4$ and $+5$ per cent (One might of course choose a different or finer regionalization, depending on the problem and the accuracy that one wishes to achieve). In each of those regions, we compute a new long-wavelength equivalent model. As an additional constraint, we impose that the individual smooth models for the different regions be close to the long-wavelength equivalent model for the average continental crust. This ensures that there are no abrupt lateral parameter changes in the complete model. Some of the technical details of the long-wavelength implementation are also provided in Table 1. The smooth model comprises 20 times less gridpoints than the original model that honours all discontinuities. A factor of, approximately, 10 lies between the storage requirements for the two models (The long-wavelength equivalent model is anisotropic with six parameters and the original model is isotropic with only three parameters). Taking into account that the time step in the smooth model is 0.09 s instead of 0.03 s in the original model, the total reduction of the computational costs amounts to a factor of roughly 50. The price to pay is a larger minimum period; 15 s instead of 5 s.

A collection of results from this numerical experiment are presented in Fig. 10. The top panel shows the vertical displacement component of the reference seismogram recorded at the station indicated in Fig. 9. Prior to 175 s the amplitudes are multiplied by 10 to enhance the visibility of the *P* waves. The cut-off period is 5 s, which is the minimum period allowed by the setup where all discontinuities coincide with element boundaries. In the row below we compare the reference seismogram and the seismogram for the long-wavelength equivalent model. The cut-off periods are 15, 20 and 25 s, where 15 s is the minimum period that we can achieve with the setup of the smooth model (approximately 28 km wide elements). For each of the cut-off frequencies both numerical solutions agree well in both the body and surface

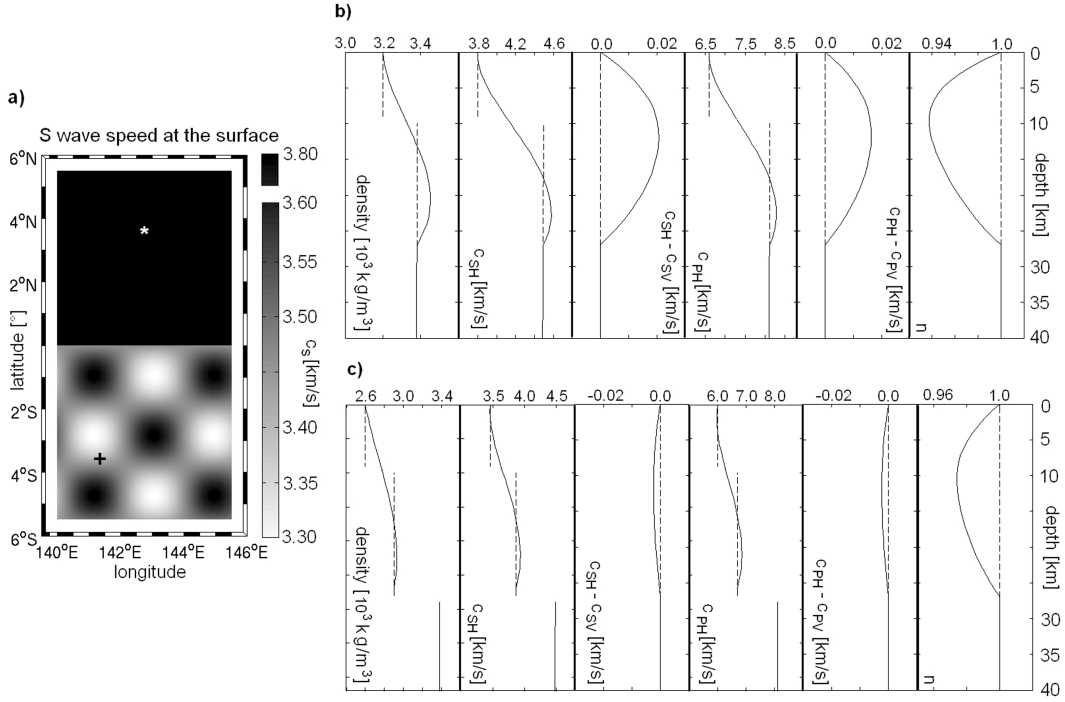

Figure 9. (a) *S*-wave speed at the surface. At northern latitudes the crust is homogeneous. At southern latitudes the *S*- and *P*-wave speeds in the crust vary by ±5 and ±3 per cent, respectively. The source location (lat $= 3.7^\circ$, lon $= 142.75^\circ$) is indicated by a star ($\star$) and the receiver location (lat $= -3.7^\circ$, lon $= 141.5^\circ$) by a plus sign (+). (b) Models of the upper 40 km for northern latitudes. The original model is plotted with dashed lines and the best long-wavelength equivalent model is plotted with solid lines. (c) Average models of the upper 40 km for southern latitudes. The original average model is plotted with dashed lines and the average of the best long-wavelength equivalent models is plotted with solid lines.

Table 1. Technical details of the numerical simulations for the model honouring the crustal discontinuities and the long-wavelength equivalent model. The values of the minimum periods are to some extent subjective but chosen conservatively.

Model	Mean width of an element (km)	No. of gridpoints	Storage required for model (GB)	Time step length (s)	No. of time steps	min. period (s)
Elements honouring discontinuities	≈9	$\approx 97 \times 10^6$	≈2.10	0.03	10000	5
Long-wavelength equivalent model	≈28	$\approx 5.2 \times 10^6$	≈0.26	0.09 s	3300	15

wave trains. This is, however, not the case when we compare the reference solution to the solution that we obtain by merely sampling the original model without implementing a long-wavelength equivalent structure. The bottom row of Fig. 10 gives an impression of the severity of this effect. As the cut-off period increases, the differences between the solutions naturally become smaller because the relative sensitivity of the surface waves to shallow structure decreases. Interestingly, the DCM also leads to more accurately modelled *P* waveforms. We defer the interpretation of the results for the 3- D model to the discussion section.

4.5.2 Short-wavelength heterogeneities

When the heterogeneities are small compared with a wavelength, we propose to first apply the DCM to the spherically symmetric part $\oplus^{1-\mathrm{D}}$ of $\oplus$ to obtain its smooth version $\oplus_s^{1-\mathrm{D}}$. The lateral heterogeneities $\delta\oplus$ are then added to obtain the smoothed 3-D model $\oplus_s = \oplus_s^{1-\mathrm{D}} + \delta\oplus$.

To demonstrate the validity of this approach, we first introduce the bandpass filtered version $\hat{\mathbf{u}}$ of the displacement field $\mathbf{u}$. The passband is assumed to be the one used for the DCM. We may therefore, write $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) \stackrel{\mathrm{DCM}}{=} \hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}})$, where the symbol $\stackrel{\mathrm{DCM}}{=}$ means equal to the extent permitted by the DCM approach. For any two displacement fields $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}})$ and $\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}})$ satisfying the wave equations

$$\rho_s\, \partial_t^2 \hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}}\big) - \nabla\cdot\mathbf{C}_s : \nabla\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}}\big) = \hat{\mathbf{f}} \tag{18}$$

and

$$\rho\, \partial_t^2 \hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) - \nabla\cdot\mathbf{C} : \nabla\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) = \hat{\mathbf{f}}; \tag{19}$$

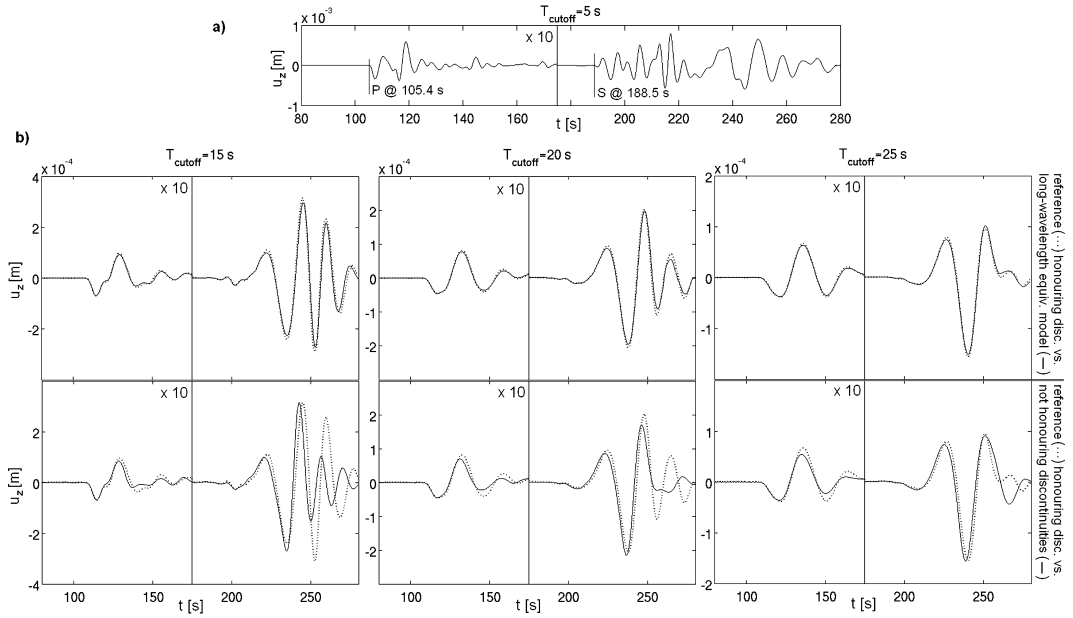

Figure 10. (a) Vertical-component of the reference displacement field recorded at the station indicated in Fig. 9. The cut-off period (low pass) is 5 s, which is also the minimum admissible period for the setup with ≈9 km wide elements. Prior to 175 s the amplitudes are multiplied by 10 so that the visibility of the *P* waves is enhanced. (b) Top row: Comparison of the reference solution (dotted line) and the solution for the long-wavelength equivalent model (solid line). Bottom row: Comparison of the reference solution (dotted line) with the solution obtained by simply sampling the original model without implementing a long-wavelength equivalent model (solid line).

we then have, by construction,

$$\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}}\big) \stackrel{\mathrm{DCM}}{=} \hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}). \tag{20}$$

In the next step we rewrite the wavefields for the heterogeneous earth models $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)$ and $\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}} + \delta\oplus)$ in the following form:

$$\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) = \hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus), \tag{21}$$

$$\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big) = \hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}}\big) + \delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big). \tag{22}$$

Hence, we need to show that $\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) \stackrel{\mathrm{DCM}}{=} \delta\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}} + \delta\oplus)$ to ensure that $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) \stackrel{\mathrm{DCM}}{=} \hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}} + \delta\oplus)$. For this, we introduce $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)$ into its corresponding wave equation:

$$(\rho^{1-\mathrm{D}} + \delta\rho)\,\partial_t^2[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)] - \nabla\cdot(\mathbf{C}^{1-\mathrm{D}} + \delta\mathbf{C}) : \nabla[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)] = \hat{\mathbf{f}}. \tag{23}$$

Eliminating a copy of the wave equation in the spherically symmetric model $\oplus^{1-\mathrm{D}}$ yields

$$\begin{aligned}&\rho^{1-\mathrm{D}}\,\partial_t^2\,\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) - \nabla\cdot\mathbf{C}^{1-\mathrm{D}} : \nabla\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)\\ &\quad = -\delta\rho\,\partial_t^2[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)] + \nabla\cdot\delta\mathbf{C} : \nabla[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)].\end{aligned} \tag{24}$$

The corresponding equation for $\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}} + \delta\oplus)$ is

$$\begin{aligned}&\rho_s^{1-\mathrm{D}}\,\partial_t^2\,\delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big) - \nabla\cdot\mathbf{C}_s^{1-\mathrm{D}} : \nabla\delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big)\\ &\quad = -\delta\rho\,\partial_t^2\big[\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big)\big] + \nabla\cdot\delta\mathbf{C} : \nabla\big[\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}}\big) + \delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big)\big].\end{aligned} \tag{25}$$

Substituting $\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}})$ by $\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}})$ in eq. (25) and using the Taylor expansion of $\delta\hat{\mathbf{u}}(\oplus_s^{1-\mathrm{D}} + \delta\oplus)$ gives

$$\begin{aligned}&\rho_s^{1-\mathrm{D}}\,\partial_t^2\,\delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big) - \nabla\cdot\mathbf{C}_s^{1-\mathrm{D}} : \nabla\delta\hat{\mathbf{u}}\big(\oplus_s^{1-\mathrm{D}} + \delta\oplus\big)\\ &\quad \stackrel{\mathrm{DCM}}{=} -\delta\rho\,\partial_t^2\big[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) + \nabla_\oplus\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)\cdot\big(\oplus_s^{1-\mathrm{D}} - \oplus^{1-\mathrm{D}}\big) + \ldots\big]\\ &\quad + \nabla\cdot\delta\mathbf{C} : \nabla\big[\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}}) + \delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus) + \nabla_\oplus\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)\cdot\big(\oplus_s^{1-\mathrm{D}} - \oplus^{1-\mathrm{D}}\big) + \ldots\big].\end{aligned} \tag{26}$$

The terms of order 2 and higher in eq. (26), symbolically written as, $\nabla_\oplus\delta\hat{\mathbf{u}}(\oplus^{1-\mathrm{D}} + \delta\oplus)\cdot(\oplus_s^{1-\mathrm{D}} - \oplus^{1-\mathrm{D}}) + \ldots$ can be neglected under the following circumstances: (1) The heterogeneities $\delta\oplus = (\delta\rho, \delta\mathbf{C})$ are small and (2) the 1-D model $\oplus^{1-\mathrm{D}}$ is close to its smoothed version $\oplus_s^{1-\mathrm{D}}$. If both conditions are satisfied, we have $\hat{\mathbf{u}}(\oplus) \stackrel{\mathrm{DCM}}{=} \hat{\mathbf{u}}(\oplus_s)$, correct to order of $\delta\oplus\cdot(\oplus_s^{1-\mathrm{D}} - \oplus^{1-\mathrm{D}})$. The DCM applied only to the spherically symmetric part $\oplus^{1-\mathrm{D}}$ of a laterally heterogeneous earth model $\oplus$ then yields a smoothed and long-wavelength equivalent model $\oplus_s = \oplus_s^{1-\mathrm{D}} + \delta\oplus$.

5 DISCUSSION AND CONCLUDING REMARKS

5.1 The advantages of the DCM technique

The principal advantage of the DCM technique is the one that motivated its development: The reduction of both the computing time and the memory requirements for numerical wave propagation. This reduction becomes particularly significant when the smallest wavelength that one wishes to model is long compared with the thickness of a crustal layer (See also the more detailed discussion below). A drastic example is the long-period surface waves. It is sufficient, in principle, for their simulation to work with a very coarse numerical grid. The presence of a thin crust, however, forces the grid spacing to be locally very fine, which leads to strongly increased numerical costs. They can be avoided using the DCM technique.

In contrast to the averaging proposed by Backus (1962), the DCM does not rely on the assumption that the thickness of the layers is orders of magnitude smaller than the smallest wavelength. In fact, the smallest wavelength in our examples (approximately 60 km) is comparable to the thickness of the crustal layers (approximately 20 km). Nevertheless, one might expect that Backus' averaging technique can yield good results for body waves because his analysis assumes a full-space and because body waves traverse thin layers quickly.

Moreover, the DCM comes—by construction—with a measure of the appropriateness of the slowly varying crustal structure, namely the dispersion curve misfit. This gives valuable information on the quality of the numerical solutions without actually computing them—assuming, of course, that the numerical algorithm does yield correct results in the case of a perfectly implemented earth model.

Finally, the DCM sheds some light onto the non-uniqueness of inverse problems that are based on an incomplete data set, that is, waveforms in a limited period range. The SA applied to the dispersion curve misfits for different models provides us with a range of models that would be acceptable solutions of an inverse problem in which the dispersion curves for the original model serve as data. The variety of models also indicates to what extent certain parameters are constrained and to what extent anisotropy is really required in the smoothed models.

5.2 Reductions of computing time and memory requirements

The amount of computing time and memory requirements that one can potentially save by replacing thin layers with a smoothed and long-wavelength equivalent structure depend strongly on both the numerical method and the earth model. We will therefore, provide only some crude estimates. First, we consider a purely isotropic grid, that is, a numerical grid where the spacing between two adjacent gridpoints, denoted by h, is equal in all coordinate directions. This is typical for most finite-difference methods but also for the spectral-element method that we based our examples on. We assume that h is chosen to be sufficiently small to permit the modelling of elastic waves with a certain minimum wavelength $\lambda_{\min}$. The memory requirements M_{c} are proportional to h^{-3} and the computing time T_{c} is proportional to h^{-4}, due to the stability condition. When thin crustal layers force us to reduce the grid spacing to $\tilde{h} < h$, then we find that T_{c} and M_{c} increase at least by a factor of $R_{\mathrm{T}} = (h/\tilde{h})^4$ and $R_{\mathrm{M}} = (h/\tilde{h})^3$, respectively. Conversely, one can say that our method of smoothing thin layers allows us to increase the grid spacing from $\tilde{h}$ to h so that the computing time and memory requirements are reduced by the factors $R_{\mathrm{T}} = (h/\tilde{h})^4$ and $R_{\mathrm{M}} = (h/\tilde{h})^3$, respectively. In more complicated cases with less symmetric grids, R_{T} and R_{M} are harder to estimate. If we assume that the vertical grid spacing h_z can be locally different from the horizontal grid spacings, then we find $R_{\mathrm{M}} = p\,(h_z/\tilde{h}_z)^1$, where p is the fraction of the model volume where the horizontal grid spacing would have to be reduced due to the presence of thin layers. For the reduction of the computing time we have, $p\,(h_z/\tilde{h}_z)^1 \leq R_T \leq p\,(h_z/\tilde{h}_z)^2$, depending on whether a reduction of the vertical grid spacing from h_z to $\tilde{h}_z$ would require a reduction of the minimum time step. Note that all these estimates are conservative because we disregarded factors such as increased communication costs in parallel codes.

In the 3-D model that we presented in Section 4.5, the reduction of the computational requirements can be estimated relatively accurately. The number of gridpoints is reduced from 97 Mio in the original model that honours the crustal discontinuities to approximately 5 Mio in the model with a wider grid spacing. This corresponds to a reduction of the storage requirements for the model itself: 2.1 GB versus 0.3 GB. Since the grid spacing is increased by a factor of approximately 3, we can increase the time step from 0.03 to 0.09 s. Combining all the effects, one would expect that the computing time decreases by a factor of approximately 58, given that the code scales perfectly. However, since the computing time usually increases super-linearly with the model size, the number 58 can be considered a conservative estimate. It is important to keep in mind that such statements are meaningful only in conjunction with an evalution of the numerical accuracy that one gains or looses with a particular method. Also this example represents, in a certain sense, a worst-case scenario. When methods allowing for local grid refinement are used, the reduction of the computational costs will naturally be smaller.

5.3 The limitations of the DCM technique and its range of validity

As we have demonstrated in the previous sections, the DCM is a useful technique that can help to reduce the computational requirements for wave propagation significantly. Still, it is an approximation that, of course, comes with its limitations: The most obvious one concerns the frequency content of the waves that one wishes to model. The higher the frequencies are the more difficult it will be to find slowly varying models that reproduce the dispersion curves of the original model acceptably well—given a fixed polynomial degree of the smoothed model. It will, for example, not be possible to find an equivalent smooth model when high-frequency body waves are considered. Fortunately, this

limitation is usually not severe because the modelling of higher-frequency waves automatically requires a refined numerical grid—irrespective of the earth structure.

An estimate of the range of validity of the DCM technique—in terms of the frequency content of the seismograms—can be obtained by determining the frequency range where the dispersion curves match sufficiently well. The term sufficiently well is naturally subjective and depends on the numerical accuracy that one wishes to achieve. In our initial example, we chose a period band 17–67 s, where the lower limit was determined by the shortes periods present in the seismograms. The upper limit was chosen such that longer periods did not have a significant influence on the smooth-model construction. The well-matching dispersion curves (see Fig. 6) served as an *a priori* indicator of the numerical accuracy of the actual seismograms. Conversely, dispersion curves that do not match in the desired frequency band would clearly suggest that the limitations of the DCM technique have been exceeded.

Sources located inside the smoothed part of the earth model are potentially problematic because the corresponding moment tensor, in principle, needs modification. The error introduced by not doing so can be minimized by designing the smooth model $\oplus_s$ so that it is close to the original one. Zero-order corrections to the moment tensor in the smoothed model are given in the paper by Capdeville & Marigo (2007).

Amplitudes of elastic waves do not enter the construction process of the slowly varying earth model. There are three reasons for this: (1) Their calculation is much more involved than the calculation of dispersion curves. The efficiency of an algorithm that incorporates amplitude information would, therefore, be inferior to the efficiency of the DCM method. (2) Amplitudes are rarely used in seismological studies. Therefore, we concentrated on the phase information. (3) Even though we do not take amplitudes into account explicitly, they are well-modelled with the smoothed crustal structures. The reason for this is that phases and amplitudes are sufficiently dependent. One constrains the other to some extent—at least in our examples. Still, it can not be excluded that there are counter-examples. In such a case, one would indeed have to incorporate amplitude information into the smooth model construction at the expense of higher computational costs.

The heart of the DCM technique is the construction of slowly varying, spherically symmetric earth models. Matching dispersion curves for laterally heterogeneous models would be technically impractical. Therefore, in the case where the lateral heterogeneities are small compared with the dominant wavelength, we suggested to apply the DCM to the spherically symmetric part $\oplus^{1-\mathrm{D}}$ of a laterally heterogeneous model $\oplus = \oplus^{1-\mathrm{D}} + \delta\oplus$ to construct a smoothed model $\oplus_s^{1-\mathrm{D}}$. The heterogeneities $\delta\oplus$ are then added to the spherically symmetric $\oplus_s^{1-\mathrm{D}}$ to give $\oplus_s$. Then we demonstrated that the model $\oplus_s$ so constructed, will lead to satisfactory results when $\oplus_s^{1-\mathrm{D}}$ is close to $\oplus^{1-\mathrm{D}}$ or when the lateral heterogeneities $\delta\oplus$ are small. The first requirement can be satisfied by construction, possibly at the expense of a less perfect dispersion curve matching. The second requirement is equivalent to saying that the heterogeneities act as scatterers. When this is not true, the DCM would have to be applied regionally. Note, however, that even if one of the requirements is not met, the method is still of first order—either in $||\oplus_s^{1-\mathrm{D}} - \oplus^{1-\mathrm{D}}||$ or in $||\delta\oplus||$. When the heterogeneities are wide compared with the dominant wavelength, the DCM can be applied regionally. We illustrated this approach with an example in Section 4.5. At least from a theoretical point of view there are complications when the wavelength of the lateral variations are comparable to the dominant wavelength of the displacement field. One then needs to evaluate carefully which one of the two approaches is preferable.

In any application the advantages and limitations of the DCM technique have to be weighted carefully. When comparing real data with synthetics one has to ensure that the remaining inaccuracies—even after the application of the DCM—are significantly smaller than the physical misfit between data and synthetics.

5.4 The 3-D model from Section 4.5

The model presented in Section 4.5 is deliberately simplistic so that it should, in principle, be reproducible with any numerical method. The geographical extension of the model is comparatively small. This is because the computation of the reference solution for a setup honouring the discontinuities is extremely costly. Limited computational resources are also the reason for which we did not yet consider an ocean layer, even though this would be of great interest and relevance.

Care should be taken when it comes to the interpretation of numerical results of this type. What we demonstrated is that the DCM applied to a 3-D problem yields satisfactory results under given circumstances. This does not mean that the DCM performs equally well, no matter what the circumstances are. The numerical accuracy achievable with the DCM is, of course, not comparable with the accuracy that can be achieved by correctly meshing all discontinuities and 3-D structures (e.g. Komatitsch & Vilotte 1998; Dumbser *et al.* 2007) and by accepting much higher computational costs. Still, the DCM can make the difference between usable and unusable synthetics.

ACKNOWLEDGMENTS

The authors want to thank Yann Capdeville for inspiring this work and for providing very insightful cristicism. Fruitful discussions with Brian Kennett, Peter Moczo, Miriam Kristeková and Jozef Kristek strongly influenced the manuscript. AF also would like to thank the seismology groups at the IPG (Paris) and the RSES (Canberra) where most of this work was done. This manuscript in particular and the author's work in general benefited greatly from the infrastructure and academic environment created by the SPICE project. Not a single numerical computation would have been possible without the technical support and patience of Jens Oeser.

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* University Science Books, Sausalito, CA.

Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.,* **67**(11), 4427–4440.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS Trans. Am. geophys. Un.,* **81**, F897.

Capdeville, Y. & Marigo, J.-J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophys. J. Int.,* **170**, 823–838.

Capdeville, Y., 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.

Chaljub, E. & Valette, B., 2004. Spectral element modelling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.,* **158,** 131–141.

Chaljub, E., Capdeville, Y. & Vilotte, J.P., 2003. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids, *J. Comp. Phys.,* **187**(2), 457–491.

Dumbser, M., Käser, M. & de la Puente, J., 2007. Arbitrary high-order finite volume schemes for seismic wave propagation on unstructured meshes in 2D and 3D, *Geophys. J. Int.,* **171**(2), 665–694.

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

Faccioli, E., Maggio, F., Paolucci, R. & Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *J. Seismol.,* **1,** 237–251.

Friederich, W. & Dalkolmo, J., 1995. Complete synthetic seismograms for a spherically symmetric Earth by a numerical computation of the Green's function in the frequency domain, *Geophys. J. Int.,* **122,** 537–550.

Igel, H., Nissen-Meyer, T. & Jahnke, G., 2002. Wave propagation in 3D spherical sections: effects of subduction zones, *Phys. Earth planet. Int.,* **132**(1–3), 219–234.

Kelly, K.R., Ward, R.W., Treitel, S. & Alford, R.M., 1976. Synthetic seismograms: a finite-difference approach, *Geophysics,* **41**(1), 2–27.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P., 1983. Optimization by simulated annealing, *Science,* **220,** 671–680.

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. seism. Soc. Am.,* **88**(2), 368–392.

Kristeková, M., Kristek, J., Moczo, P. & Day, S.M., 2006. Misfit criteria for quantitative comparison of seismograms, *Bull. seism. Soc. Am.,* **96**(5), 1836–1850.

Love, A.E.H., 1944. *Mathematical Theory of Elasticity,* Dover Publications, New York.

Moczo, P., Kristek, J., Vavrycuk, V., Archuleta, R.J. & Halada, L., 2002. 3D heterogeneous staggered-grid finite-difference modelling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities, *Bull. seism. Soc. Am.,* **92**(8), 3042–3066.

Priolo, E., Carcione, J.M. & Seriani, G., 1994. Numerical simulation of interface waves by high-order spectral modeling techniques, *J. acoust. Soc. Am.,* **95**(2), 681–693.

Schoenberg, M. & Muir, F., 1989. A calculus for finely layered anisotropic media, *Geophysics,* **54**(5), 581–589.

Seriani, G., 1998. 3-D large-scale wave propagation modeling by spectral element method on Cray T3E multiprocessor, *Comp. Methods Appl. Mech. Engrg.,* **164**(1–2), 235–247.

Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods, *Geophys. J. Int.,* **168**(3), 1153–1174.

Strang, G. & Nguyen, T., 1996. *Wavelets and Filter Banks,* Wellesley-Cambridge Press, Wellesley, MA.

Stuart, A. & Pavliotis, G., 2007. *Multiscale Methods: Averaging and Homogenization,* Springer, Berlin.

Takeuchi, H. & Saito, M., 1972. Seismic surface waves, in *Methods in Computational Physics,* Vol. 11, pp. 217–295, Academic Press, New York.

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