Geophys. J. Int. (2003) **155,** 289–307

Surface wave higher-mode phase velocity measurements using a roller-coaster-type algorithm

Éric Beucler,* Éléonore Stutzmann and Jean-Paul Montagner

Laboratoire de sismologie globale, IPGP, 4 *place Jussieu,* 75252 *Paris Cedex* 05, *France. E-mail: beucler@ipgp.jussieu.fr*

Accepted 2003 May 20. Received 2003 January 6; in original form 2002 March 14

SUMMARY

In order to solve a highly non-linear problem by introducing the smallest *a priori* information, we present a new inverse technique called the 'roller coaster' technique and apply it to measure surface wave mode-branch phase velocities. The fundamental mode and the first six overtone parameter vectors, defined over their own significant frequency ranges, are smoothed average phase velocity perturbations along the great circle epicentre–station path. These measurements explain well both Rayleigh and Love waveforms, within a maximum period range included between 40 and 500 s. The main idea of this technique is to first determine all possible configurations of the parameter vector, imposing large-scale correlations over the model space, and secondly to explore each of them locally in order to match the short-wavelength variations. The final solution which achieves the minimum misfit of all local optimizations, in the least-squares sense, is then hardly influenced by the reference model. Each mode-branch *a posteriori* reliability estimate turns out to be a very powerful instrument in assessing the phase velocity measurements. Our Rayleigh results for the Vanuatu–California path seem to agree correctly with previous ones.

Key words: inverse problem, seismic tomography, surface waves, waveform analysis.

1 INTRODUCTION

Over the last two decades, the resolution of global tomographic models has been greatly improved, because of the increase in the amount and the quality of data, and due to more and more sophisticated data processing and inversion schemes (Woodhouse & Dziewonski 1984, 1986; Montagner 1986; Nataf *et al.* 1986; Giardini *et al.* 1987; Montagner & Tanimoto 1990; Tanimoto 1990; Zhang & Tanimoto 1991; Su *et al.* 1994; Li & Romanowicz 1995; Romanowicz 1995; Trampert & Woodhouse 1995; Laske & Masters 1996; Ekström *et al.* 1997; Grand *et al.* 1997; van der Hilst *et al.* 1997; Liu & Dziewonski 1998; Ekström & Dziewonski 1998; Laske & Masters 1998; Mégnin & Romanowicz 2000; Ritsema & van Heijst 2000, among others). These models are derived from surface wave phase velocities and/or body wave traveltimes (or waveforms) and/or free-oscillation splitting measurements. Body wave studies provide high-resolution models but suffer from the inhomogeneous distribution of earthquakes and recording stations, even when considering reflected or diffracted phases. On the other hand, the surface wave fundamental mode is mainly sensitive to the physical properties of the upper mantle. So, the investigation of the transition zone on a global scale, which plays a key role in mantle convection, can only be achieved by using higher-mode surface waves.

A first attempt at providing a global tomographic model using these waves has been proposed by Stutzmann & Montagner (1994), but with a limited amount of data. More recently, van Heijst & Woodhouse (1999) computed degree-12 phase velocity maps of the fundamental mode and the four first overtones for both Love and Rayleigh waves. These data have been combined with body wave traveltimes measurements and free-oscillation splitting measurements, to provide a global tomographic model with a high and uniform resolution over the whole mantle (Ritsema *et al.* 1999; van Heijst *et al.* 1999). The most recent S_H model for the whole mantle was proposed by Mégnin & Romanowicz (2000). This degree-24 model results from waveform inversion of body and surface Love waves, including fundamental and higher modes and introducing cross-branch coupling.

Extracting information from higher-mode surface waves is a difficult task. The simultaneous arrivals (Fig. 3 in Section 3) and the interference between the different mode-branches make the problem very underdetermined and non-linear. To remove the non-linearity, Cara & Lévêque (1987) and Lévêque *et al.* (1991) compute the cross-correlogram between the data and monomode synthetic seismograms and

*Now at: École Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France.

invert the amplitude and the phase of the filtered cross-correlogram. On the other hand, Nolet *et al.* (1986) and Nolet (1990) use an iterative inverse algorithm to fit the waveform in the time domain and increase the model complexity within the iterations. These two methods provide directly a 1-D model corresponding to an average epicentre–station path. They were first used 'manually', which limited the amount of data that could be processed. The exponential increase in the amount of good-quality broad-band data has made necessary the automation of most parts of the data processing and an automatic version of these methods has been proposed by Debayle (1999) for the waveform inversion technique of Cara & Lévêque (1987) and by Lebedev (2000) and Lebedev & Nolet (2003) for the partition waveform inversion.

Stutzmann & Montagner (1993) split the inversion into two steps; at each iteration, a least-squares optimization to measure phase velocities is followed by an inversion to determine the 1-D *S*-wave velocity model, in order to gain insight into the factors that control the depth resolution. They retrieve the phase velocity for a set of several seismograms recorded at a single station and originating from earthquakes located in the same area in order to improve the resolution. Another approach has been followed by van Heijst & Woodhouse (1997) who proposed a mode-branch stripping technique based on monomode cross-correlation functions. Phase velocity and amplitude perturbations are determined for the most energetic mode-branch, the waveform of which is then subtracted from the seismogram in order to determine the second most energetic mode-branch phase velocity and amplitude perturbations, and so on. More recently, Yoshizawa & Kennett (2002) used the neighbourhood algorithm (Sambridge 1999a,b) to explore the model space in detail and to obtain directly a 1-D velocity model which achieves the minimum misfit. It is difficult to compare the efficiency of these methods because they all follow different approaches to taking account of the non-linearity of the problem. Up to now, it has only been possible to compare tomographic results obtained using these different techniques.

In this paper, we introduce a new semi-automatic inverse procedure, the 'roller coaster' technique (owing to the shape of the misfit curve displayed in Fig. 6b in Section 3.4.1), to measure fundamental and overtone phase velocities both for Rayleigh and Love waves. This method can be applied either to a single seismogram or to a set of seismograms recorded at a single station. To deal with the non-linearity of the problem, the roller coaster technique combines the detection of all possible solutions at a large scale (which means solutions of large-wavelength variations of the parameter vector over the model space), and local least-squares inversions close to each of them, in order to match small variations of the model. The purpose of this article is to present an inverse procedure that introduces as little *a priori* information as possible in a non-linear scheme. So, even using a straightforward phase perturbation theory, we show how this algorithm detects and converges towards the best global misfit model. The roller coaster technique is applied to a path average theory but can be later adapted and used with a more realistic wave propagation theory. One issue of this study is to provide a 3-D global model which does not suffer from strong *a priori* constraints during the inversion and which then can be used in the future as a reference model.

We describe hereafter the forward problem and the non-linear inverse approach developed for solving it. An essential asset of this technique is to provide quantitative *a posteriori* information, in order to assess the accuracy of the phase velocity measurements. Resolution tests on both synthetic and real data are presented for Love and Rayleigh waves.

2 FORWARD PROBLEM

Following the normal-mode summation approach, a long-period seismogram can be modelled as the sum of the fundamental mode ($n = 0$) and the first higher modes ($n \geq 1$), hereafter referred to as FM and HM, respectively. Eigenfrequencies and eigenfunctions are computed for both spheroidal and toroidal modes in a 1-D reference model, PREM (Dziewonski & Anderson 1981) in our case. Stoneley modes are removed, then the radial order n for the spheroidal modes corresponds to Okal's classification (Okal 1978). In the following, all possible sorts of coupling between toroidal and spheroidal mode-branches (Woodhouse 1980; Lognonné & Romanowicz 1990; Deuss & Woodhouse 2001) and off-great-circle propagation effects (Woodhouse & Wong 1986; Laske & Masters 1996) are neglected.

For a given recorded long-period seismogram, the corresponding synthetic seismogram is computed using the formalism defined by Woodhouse & Girnius (1982). In the most general case, the displacement **u**, corresponding of the first surface wave train, in the time domain, can be written as

$$\mathbf{u}(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{j=0}^{n} A_j(\mathbf{r}, \omega) \exp[i\Phi_j(\mathbf{r}, \omega)] \exp(i\omega t)\, d\omega, \tag{1}$$

where **r** is the source–receiver spatial position, ω is the angular frequency and where A_j and Φ_j represent the amplitude and the phase of the jth mode-branch, respectively, in the frequency domain. In the following, the recorded and the corresponding synthetic seismogram spectra (computed in PREM) are denoted by (R) and (S), respectively.

In the Fourier domain, following Kanamori & Given (1981), a recorded seismogram spectrum can be written as

$$A^{(\mathrm{R})}(\mathbf{r}, \omega) \exp\left[i\Phi^{(\mathrm{R})}(\mathbf{r}, \omega)\right] = \sum_{j=0}^{n} B_j(\mathbf{r}, \omega) \exp\left[i \left(\Psi_j(\mathbf{r}, \omega) - \frac{\omega a \Delta}{C_j^{(\mathrm{R})}(\mathbf{r}, \omega)} \right) \right], \tag{2}$$

where a is the radius of the Earth, Δ is the epicentral distance (in radians) and $C_j^{(\mathrm{R})}(\mathbf{r}, \omega)$ is the real average phase velocity along the epicentre–station path of the jth mode-branch, which we wish to measure. The term $B_j(\mathbf{r}, \omega)$ includes source amplitude and geometrical spreading, whereas $\Psi_j(\mathbf{r}, \omega)$ corresponds to the source phase. The instrumental response is included in both terms and this expression is valid for both Rayleigh and Love waves. The phase shift due to the propagation in the real medium then resides in the term $\exp\left[-i\omega a \Delta / C_j^{(\mathrm{R})}(\mathbf{r}, \omega)\right]$.

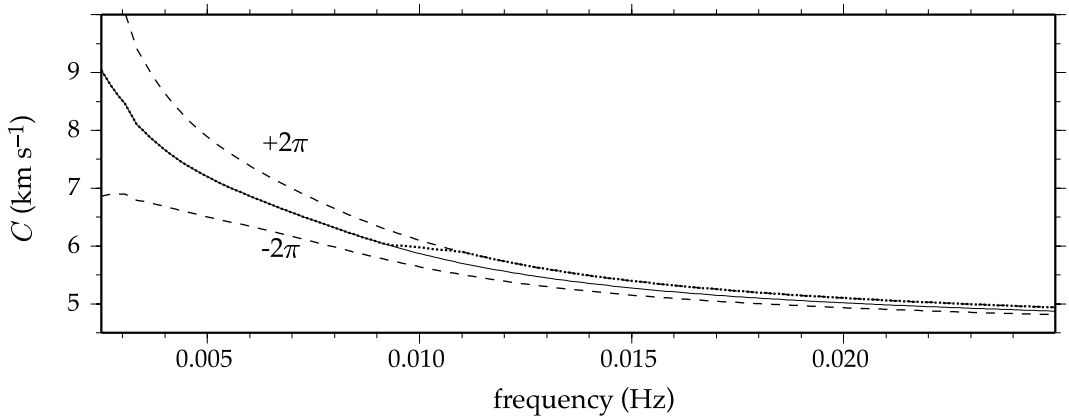

Figure 1. Illustration of possible 2π phase jumps over the whole frequency range (dashed lines) or localized around a given frequency (dotted line). The reference phase velocity used to compute these three curves is represented as a solid line.

Considering that, to first order, the effect of a phase perturbation dominates over that of the amplitude perturbation (Li & Tanimoto 1993), and writing the real slowness as a perturbation of the synthetic slowness (computed in the 1-D reference model), eq. (2) becomes

$$A^{(\mathrm{R})}(\mathbf{r},\omega)\exp\left[i\,\Phi^{(\mathrm{R})}(\mathbf{r},\omega)\right]=\sum_{j=0}^{n}A_j^{(\mathrm{S})}(\mathbf{r},\omega)\exp\left\{i\left[\Psi_j(\mathbf{r},\omega)-\frac{\omega a\Delta}{C_j^{(\mathrm{S})}(\omega)}-\chi\right]\right\}, \tag{3}$$

where

$$\chi=\omega a\Delta\left(\frac{1}{C_j^{(\mathrm{R})}(\mathbf{r},\omega)}-\frac{1}{C_j^{(\mathrm{S})}(\mathbf{r},\omega)}\right). \tag{4}$$

Let us now denote by $\mathbf{p}_j(\mathbf{r},\omega)$, the dimensionless parameter vector of the jth mode-branch defined by

$$\mathbf{p}_j(\mathbf{r},\omega)=\frac{C_j^{(\mathrm{R})}(\mathbf{r},\omega)-C_j^{(\mathrm{S})}(\omega)}{C_j^{(\mathrm{R})}(\omega)}. \tag{5}$$

Finally, introducing the synthetic phase $\Phi_j^{(\mathrm{S})}(\mathbf{r},\omega)$, as the sum of the source phase and the phase shift due to the propagation in the reference model, the forward problem can be expressed as

$$\mathbf{d}=g(\mathbf{p}),$$

$$A^{(\mathrm{R})}(\mathbf{r},\omega)\exp\left[i\,\Phi^{(\mathrm{R})}(\mathbf{r},\omega)\right]=\sum_{j=0}^{n}A_j^{(\mathrm{S})}(\mathbf{r},\omega)\exp\left\{i\left[\Phi_j^{(\mathrm{S})}(\mathbf{r},\omega)+\frac{\omega a\Delta}{C_j^{(\mathrm{S})}(\omega)}\mathbf{p}_j(\mathbf{r},\omega)\right]\right\}. \tag{6}$$

For practical reasons, the results presented in this paper are computed following a forward problem expression based on phase velocity perturbation expanded to third order (eq. A5). When considering an absolute perturbation range lower than 10 per cent, results are, however, identical to those computed following eq. (6) (see Appendix A).

Formally, eq. (6) can be summarized as a linear combination of complex cosines and sines and for this reason, a 2π undetermination remains for every solution. For a given parameter $\mathbf{p}_j(\mathbf{r},\omega)$, it is obvious that two other solutions can be found by a 2π shift such as

$$\mathbf{p}_j^{+}(\mathbf{r},\omega)=\mathbf{p}_j(\mathbf{r},\omega)+\frac{2\pi C_j^{(\mathrm{S})}(\omega)}{\omega a\Delta}\quad\text{and}\quad\mathbf{p}_j^{-}(\mathbf{r},\omega)=\mathbf{p}_j(\mathbf{r},\omega)-\frac{2\pi C_j^{(\mathrm{S})}(\omega)}{\omega a\Delta}. \tag{7}$$

As an example of this feature, all the phase velocity curves presented in Fig. 1 satisfy eq. (6). This means that 2π phase jumps can occur over the whole frequency range but can also be localized around a given frequency. Such an underdetermination as expressed in eq. (6) and such a non-unicity, in most cases due to the 2π phase jumps, are often resolved by imposing some *a priori* constraints in the inversion. *A contrario*, the roller coaster technique explores a large range of possible solutions, with the smallest *a priori* as possible, before choosing the model that achieves the minimum misfit.

3 DESCRIPTION OF THE ROLLER COASTER TECHNIQUE

The method presented in this paper is a hybrid approach, combining detection of all possible large-scale solutions (which means solutions of long-wavelength configurations of the parameter vector) and local least-squares optimizations starting from each of these solutions, in order to match the short-wavelength variations of the model space. The different stages of the roller coaster technique are presented in Fig. 2 and described hereafter. The first three stages are devoted to the reduction of the problem underdetermination, while the non-linearity and the non-unicity are taken into account in the following steps.

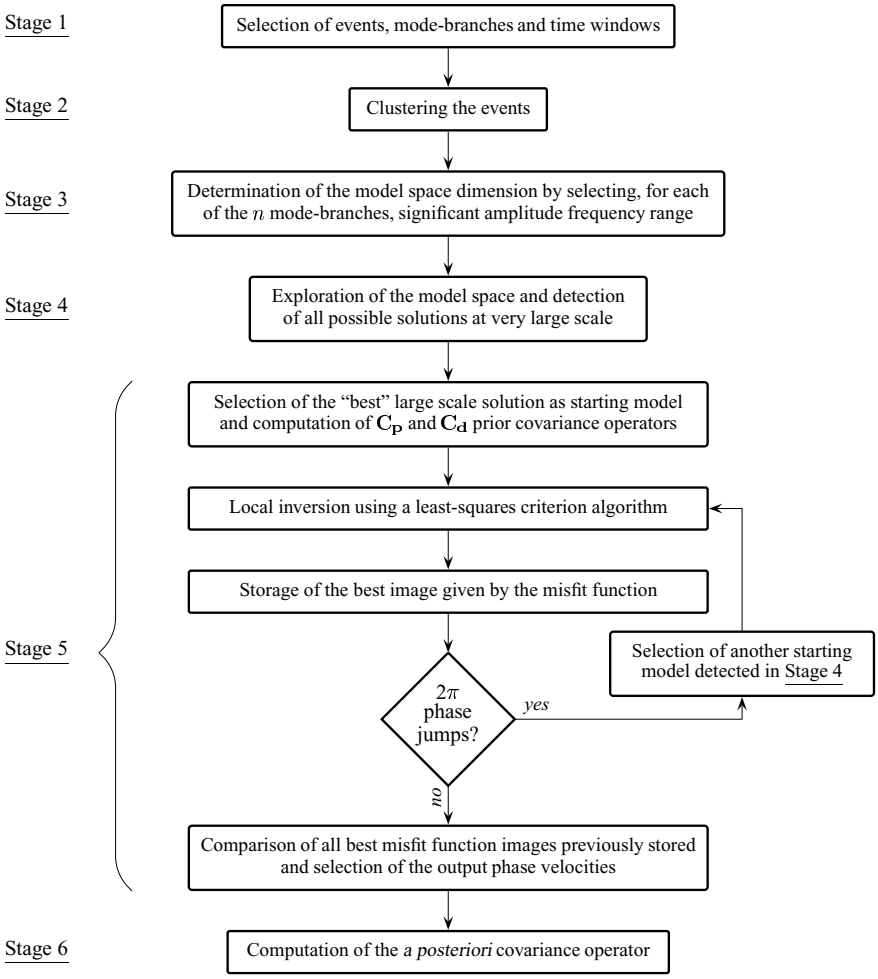

Figure 2. Schematic diagram of the roller coaster technique. See Section 3 for details.

3.1 Selection of events, mode-branches and time windows

Events with epicentral distances larger than 55° and shorter than 135° are selected. Thus, the FM is well separated in time from the HM (Fig. 3), and the first and the second surface wave trains do not overlap. Since the FM signal amplitude is much larger than the HM amplitude for about 95 per cent of earthquakes, each seismogram (real and synthetic) is temporally divided into two different time windows, corresponding to the FM and to the HM parts of the signal. An illustration of this amplitude discrepancy in the time domain is displayed in Fig. 3(b) and when focusing on Fig. 4(a), the spectrum amplitude of the whole real signal (FM+HM) is largely dominated by the FM one. Eight different pickings defining the four time windows, illustrated in Fig. 3(a), are computed using synthetic mode-branch wave trains and are checked manually. For this reason, this method is not completely automated, but this picking step is necessary to assess the data quality and the consistency between recorded and synthetic seismograms. In Appendix B, we show that the phase velocity measurements are not significantly affected by a small change in the time window dimensions. An advantage of this temporal truncation is that, whatever the amplitude of the FM, the HM part of the seismograms can always be treated. Hence, the forward problem is now split into two equations, corresponding to the FM and to the HM parts, respectively.

$$A^{(\mathrm{R})}_{\mathrm{FM}}(\mathbf{r},\omega)\exp\left[i\,\Phi^{(\mathrm{R})}_{\mathrm{FM}}(\mathbf{r},\omega)\right] = A^{(\mathrm{S})}_{0}(\mathbf{r},\omega)\exp\left\{i\left[\Phi^{(\mathrm{S})}_{0}(\mathbf{r},\omega)+\frac{\omega a\,\Delta}{C^{(\mathrm{S})}_{0}(\omega)}\mathbf{p}_0(\mathbf{r},\omega)\right]\right\} \tag{8}$$

and

$$A^{(\mathrm{R})}_{\mathrm{HM}}(\mathbf{r},\omega)\exp\left[i\,\Phi^{(\mathrm{R})}_{\mathrm{HM}}(\mathbf{r},\omega)\right] = \sum_{j=1}^{6} A^{(\mathrm{S})}_{j}(\mathbf{r},\omega)\exp\left\{i\left[\Phi^{(\mathrm{S})}_{j}(\mathbf{r},\omega)+\frac{\omega a\,\Delta}{C^{(\mathrm{S})}_{j}(\omega)}\mathbf{p}_j(\mathbf{r},\omega)\right]\right\}. \tag{9}$$

Seismograms (real and synthetic) are bandpass filtered between 40 and 500 s. In this frequency range, only the first six overtone phase velocities can be efficiently retrieved. Tests on synthetic seismograms (up to $n = 15$) with various depths and source parameters have shown that the HM for $n \geq 7$ have negligible amplitudes in the selected time and frequency windows.

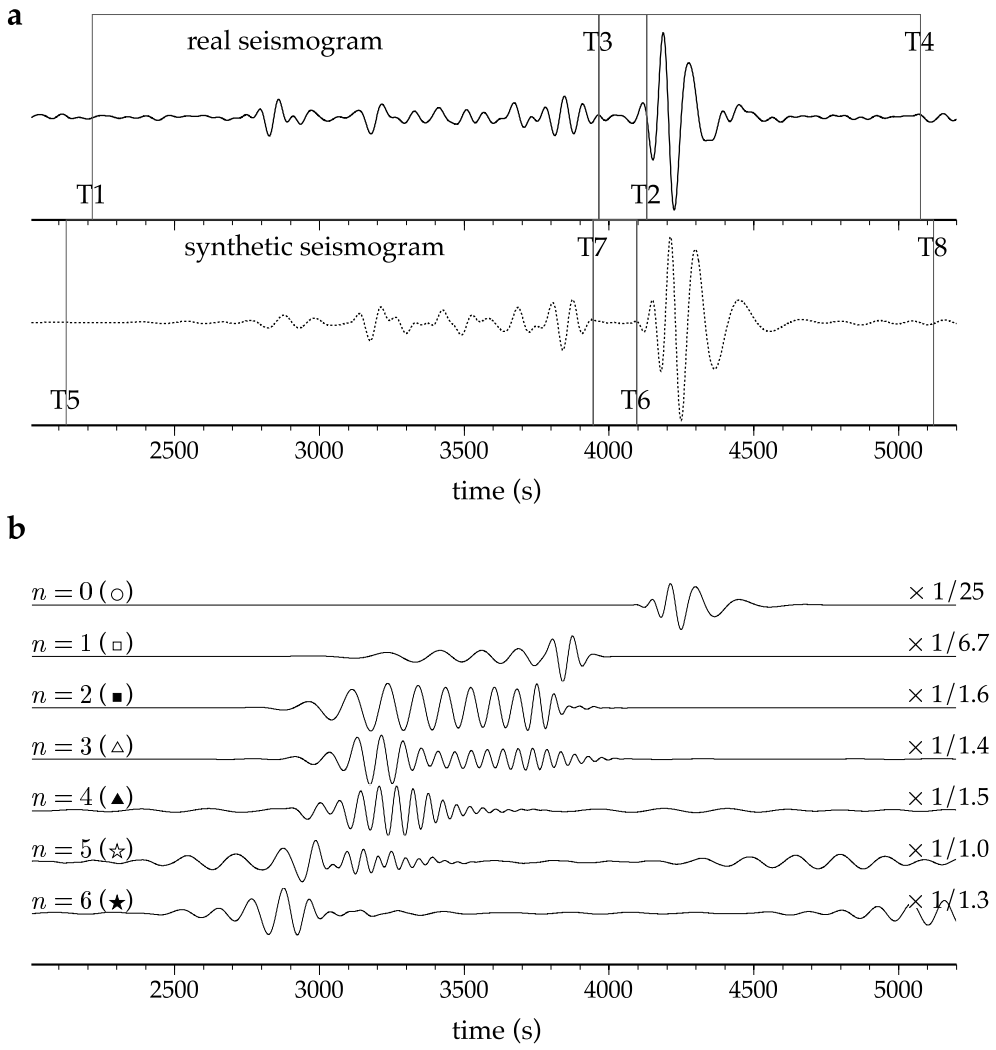

Figure 3. (a) Real vertical seismogram (solid line) and its corresponding synthetic computed in PREM (dotted line). The earthquake underlying this waveform occurred on 1993 September 4 in Afghanistan (36°N, 70°E, depth of 190 km) and was recorded at the CAN GEOSCOPE station (Australia). The epicentral distance is estimated at around 11 340 km. Both waveforms are divided into two time windows corresponding to the higher modes (T1–T2, T5–T6) and to the fundamental mode (T3–T4, T7–T8). (b) The contribution of each synthetic monomode shows the large-amplitude discrepancy and time delay between the fundamental mode and the overtones. The different symbols refer to the spectra displayed in Fig. 4.

3.2 Clustering the events

Following eq. (8), a single seismogram is sufficient to measure the FM phase velocity, whereas for the HM (eq. 9) the problem is still highly underdetermined since the different HM group velocities are very close. This can be avoided by a reduction of the number of independent parameters considering mathematical relations between different mode-branch phase velocities. The consequence of such an approach is to impose a strong *a priori* knowledge on the model space, which may be physically unjustified. Another way to reduce this underdetermination is to increase the amount of independent data while keeping the parameter space dimension constant. Therefore, all sufficiently close events are clustered into small areas, and each individual ray path belonging to the same box is considered to give equivalent results as a common ray path. This latter approach was followed by Stutzmann & Montagner (1993), but with 5×5 deg^2 boxes independently of epicentral distance and azimuth values, due to the limited number of data. Here, in order to prevent any bias induced by the clustering of events too far away from one to another, and to be consistent with the smallest wavelength, boxes are computed with a maximum aperture angle of 2° and 4° in the transverse and longitudinal directions, respectively (Fig. 5), with respect to the great circle path. The boxes are computed in order to take into account as many different depths and source mechanisms as possible. The FM phase velocity inversion is performed for each path between a station and a box, whereas the HM phase velocities are only measured for the boxes including three or more events. Since only the six first mode-branches spectra are inverted, the maximum number of events per box is set to eight. The use of different events implies average phase velocity measurements along the common ray paths which can be unsuitable for short epicentral distances, but increases the accuracy of the results for the epicentral distances considered.

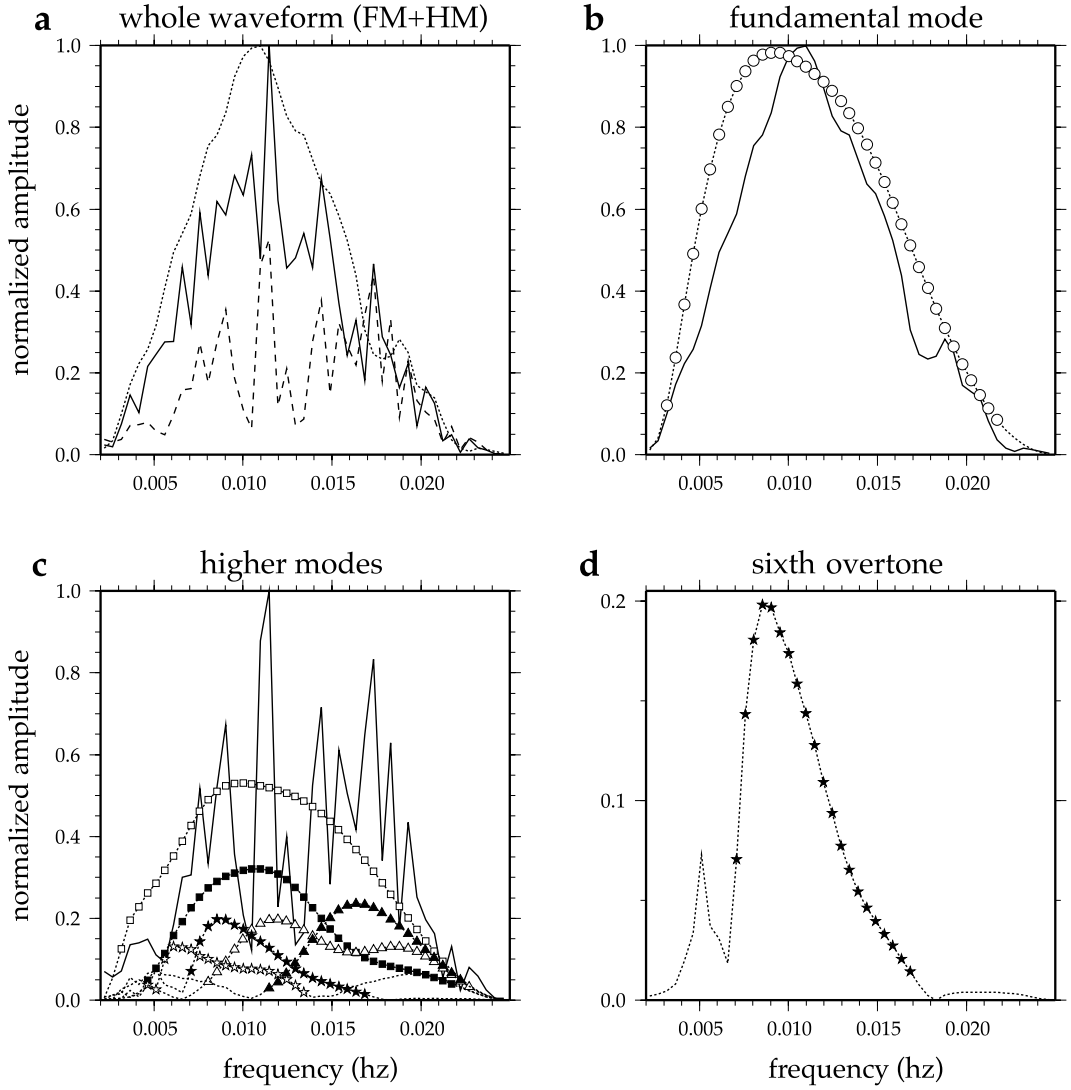

Figure 4. (a) The normalized amplitude spectra of the whole real waveform (solid line) displayed in Fig. 3(a). The real FM part of the signal (truncated between T3 and T4) is represented as a dotted line and the real HM part (between T1 and T2) as a dashed line. (b). The solid line corresponds to the normalized spectrum amplitude of the real signal truncated between T3 and T4 (Fig. 3a). The corresponding synthetic FM is represented as a dotted line and only the frequency range represented by the white circles is selected as being significant. (c) Selection of HM inversion frequency ranges using synthetic significant amplitudes. The solid line corresponds to the real HM signal, picked between T1 and T2 (Fig. 3a). For each mode-branch (dotted lines), only the frequency ranges defined by the symbols (according to Fig. 3b) are retained for the inversion. (d) Close up of the sixth synthetic overtone, in order to visualize the presence of lobes and the weak contribution frequency range in the spectrum amplitude. The stars delimit the selected frequency range.

3.3 Determination of the model space dimension

Real and synthetic amplitude spectra are normalized in order to minimize the effects due to the imprecision of source parameters and of instrumental response determination. As presented in Fig. 4, a synthetic mode-branch spectrum is frequently composed by several lobes due to the source mechanism. Between each lobe and also near the frequency range edges due to the bandpass filter, the amplitude strongly decreases down to zero, and therefore phase velocities are absolutely not constrained at these frequencies. It is around these frequencies that possible local 2π phase jumps may occur (Fig. 1). Then, we decide to reduce the model space dimension in order to take into account only well-constrained points. For each spectrum, the selection of significant amplitudes, with a threshold fixed to 10 per cent of the mean maximum spectra amplitude, defines the inverted frequency range. In the case of several lobes in a synthetic mode-branch amplitude spectrum, only the most energetic one is selected as shown in Figs 4(c) and (d). For a given mode-branch, the simultaneous use of different earthquakes implies a discrimination criterion based upon a mean amplitude spectrum of all spectra, which tends to increase the dimensions of the significant frequency range. The normalization and this selection of each mode-branch significant amplitudes is also a way to include surface wave radiation pattern information in the procedure. Changes in source parameters can result in changes in the positions of the lobes in the mode-branch amplitude spectra over the whole frequency range (40–500 s). In the future, it will be essential to include these possible biases in the scheme and then to simultaneously invert moment tensor, location and depth.

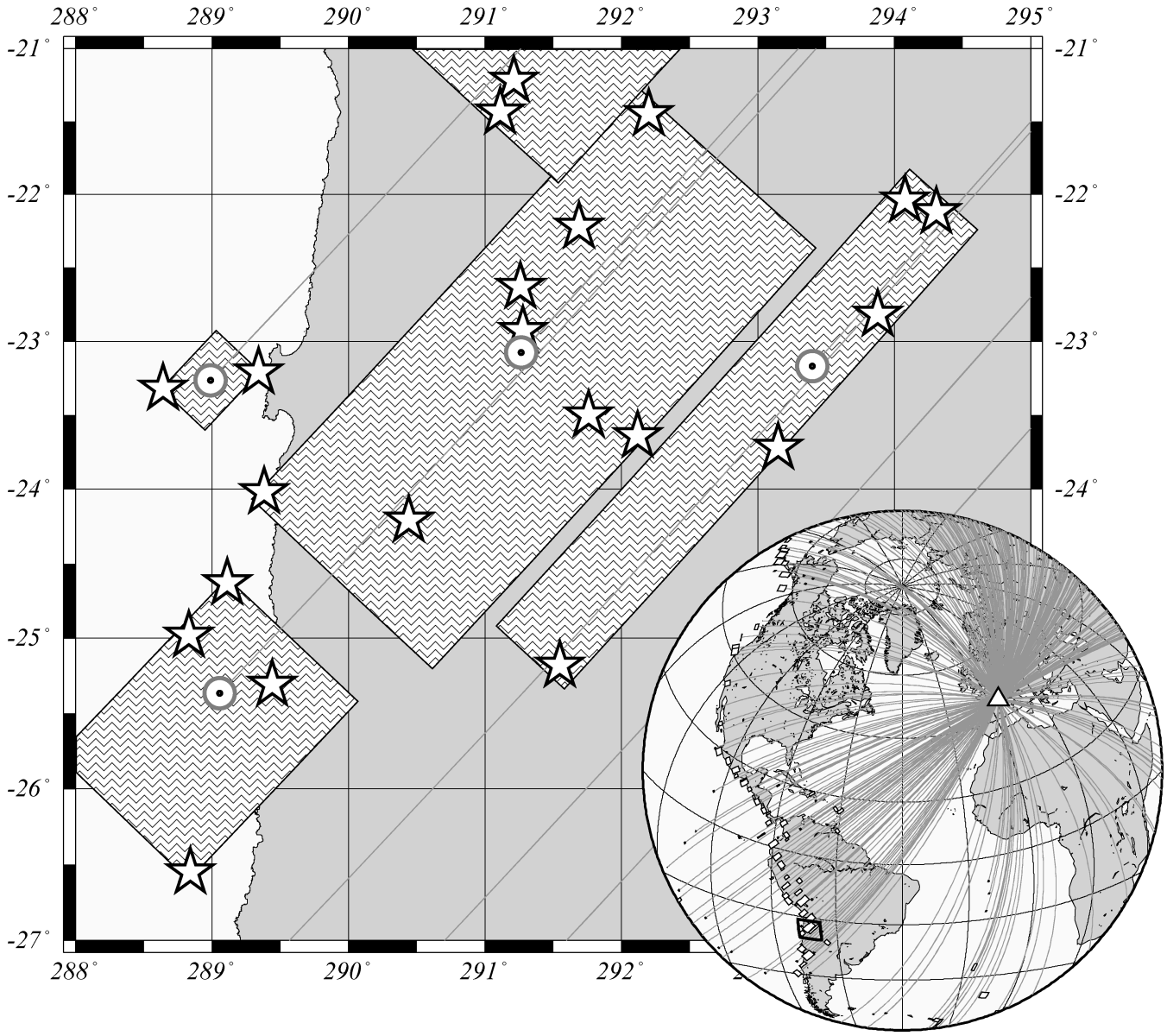

Figure 5. Geographical distribution of inversion boxes for the SSB GEOSCOPE station case. The enlarged area is defined by the bold square in the inset (South America). Black stars denote epicentres and hatched grey boxes join each inversion group. Each common ray path (grey lines) starts from the barycentre (circles) of all events belonging to the same box. The maximum number of seismograms per box is fixed at eight.

3.4 Exploration of the model space at very large scale

The main idea of this stage is to test a large number of phase velocity large-scale perturbations with the view of selecting several starting vectors for local inversions (see Section 3.5). The high non-linearity of the problem is mainly due to the possible 2π phase jumps. And, even though the previous stage (see Section 3.3) prevents the shifts inside a given mode-branch phase velocity curve, 2π phase jumps over the whole selected frequency range are still possible. For this reason a classical gradient least-squares optimization (Tarantola & Valette 1982a) is inadequate. In a highly non-linear problem, a least-squares inversion only converges towards the best misfit model that is closest to the starting model and the number of iterations cannot change this feature. On the other hand, a complete exploration of all possible configurations in the parameter space is still incompatible with a short computation time procedure. Therefore, an exploration of the model space is performed at very large scale, in order to detect all possible models that globally explain the data set well.

3.4.1 Fundamental mode case

When considering a single mode-branch, the number of parameter vector components is rather small. The FM large-scale exploration can then be more detailed than in the HM case. Considering that, at low frequencies, data are correctly explained by the 1-D reference model, the

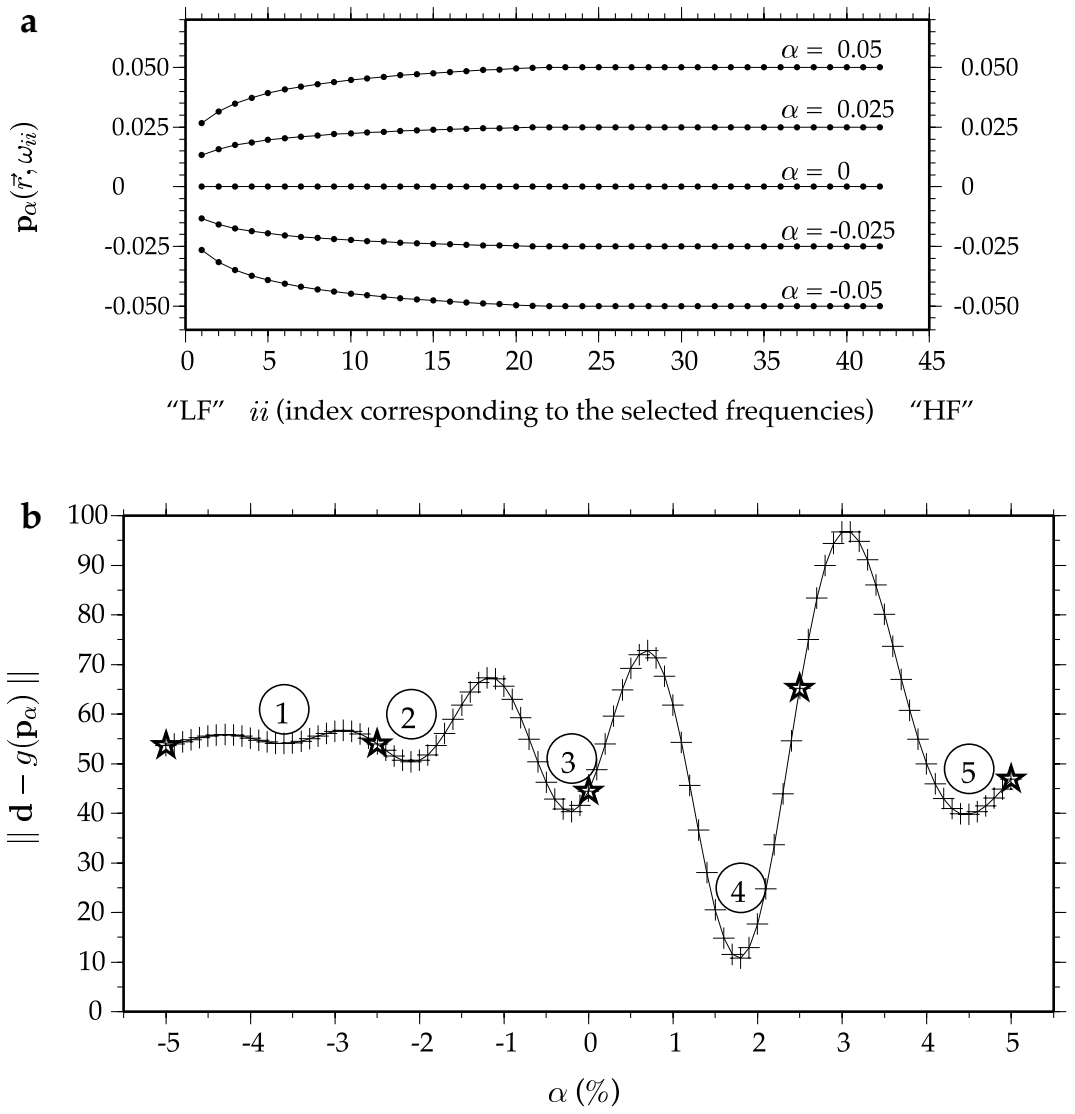

Figure 6. (a) Five examples of the FM parameter vector configurations during the exploration of the model space at large scale corresponding to α values equal to -5, $-$, 0, $+2.5$ and $+5$ per cent. The selected points for which the phase velocity is measured (see Section 3.3) are ordered into parameter vector components according to increasing frequency values. The first indices then correspond to the low-frequency components (LF) and the last ones to the high-frequency (HF) components. Varying the exploration factor α, different perturbation shapes are then modelled and the misfit between data and the image of the corresponding vector is measured (represented in the figure below). (b) The misfit in the FM case, symbolized by $+$, is the expression of the difference between data and the image of the tested model (referred to as $\mathbf{p}_\alpha$) through the g function (eq. 8). The α values are expressed as a percentage with respect to the PREM. As an example, the five stars correspond to the misfit values of the five models represented in the figure above. The circles represent the best α values and the corresponding vectors are then considered as possible starting models for the next stage.

dimensionless phase velocity perturbation (referred to as $\mathbf{p}_\alpha$) can be modelled as shown in the five examples displayed in Fig. 6(a). Basically, the low-frequency component perturbations are smaller than the high-frequency ones. However, if such an assumption cannot be made, the simplest way to explore the model space is then by fixing an equal α perturbation value for all the components. The main idea is to impose strong correlations between all the components in order to estimate how high the non-linearity is.

Varying α enables one to compute different parameter vectors and solving eq. (8) to measure the distance between data and the image of a given model through the g function, integrated over the whole selected frequency range. Considering that only small perturbations can be retrieved, the exploration range is limited between -5 and $+5$ per cent, using an increment step of 0.1 per cent. The result of such an exploration is displayed in Fig. 6(b) and clearly illustrates the high non-linearity and non-unicity of the problem. In a weakly non-linear problem, the misfit curve (referred to as $||\mathbf{d} - g(\mathbf{p}_\alpha)||$) should exhibit only one minimum. This would indicate that, whatever the value of the starting model, a gradient algorithm always converges towards the same final model, the solution is then unique. In our case, Fig. 6(b) shows that, when choosing the reference model (i.e. $\alpha = 0$ per cent) as the starting model, a gradient least-squares optimization converges to the nearest best-fitting solution (corresponding to the third circle), and could never reach the global best-fitting model (in this example represented

by the fourth circle). Therefore, in order not to *a priori* limit the inversion result around a given model, all minima of the misfit curve (Fig. 6b) are detected and the corresponding vectors are considered as possible starting models for local optimizations (see Section 3.5).

3.4.2 Higher-mode case

The introduction of several mode-branches simultaneously is much more difficult to treat and it becomes rapidly infeasible to explore the model space as accurately as performed for the FM. However, a similar approach is followed. In order to preserve a low computation time procedure, the increment step of α is fixed at 1 per cent. The different parameter vectors are computed as previously explained in Section 3.4.1 (the shape of each mode-branch subvector is the same as the examples displayed in Fig. 6a). In order to take into account any possible influence of one mode-branch on another, all combinations are tested systematically. Three different explorations of the model space are performed within three different research ranges: [−4.5 to +1.5 per cent], [−3 to +3 per cent] and [−1.5 to +4.5 per cent]. For each of them, 7^6 possibilities of the parameter vector are modelled and the misfit between data and the image of the tested vector through the g function is computed. This approach is almost equivalent to performing a complete exploration in the range [−4.5 to +4.5 per cent], using a step of 0.5 per cent, but less time consuming. Finally, all misfit curve minima are detected and, according to a state of null information concerning relations between each mode-branch phase velocities, all the corresponding vectors are retained as possible starting models. Thus, any association between each starting model subvectors is allowed.

3.5 Matching the short-wavelength variations of the model

In this section, algorithms, notation and comments are identical for both FM and HM. Only the main ideas of the least-squares criterion are outlined. A complete description of this approach is given by Tarantola & Valette (1982a,b) and by Tarantola (1987). Some typical features related to the frequency/period duality are also detailed.

3.5.1 The gradient least-squares algorithm

The main assumption which leads us to use such an optimization is to consider that starting from the large-scale parameter vector (see Section 3.4), the non-linearity of the problem is largely reduced. Hence, to infer the model space from the data space, a gradient least-squares algorithm is performed (Tarantola & Valette 1982a). The expression of the model (or parameter) at the kth iteration is given by

$$\mathbf{p}_k = \mathbf{p}_0 + \mathbf{C_p} \cdot \mathbf{G}_{k-1}^{\mathrm{T}} \cdot \left[\mathbf{C_d} + \mathbf{G}_{k-1} \cdot \mathbf{C_p} \cdot \mathbf{G}_{k-1}^{\mathrm{T}}\right]^{-1} \cdot \left[\mathbf{d} - g(\mathbf{p}_{k-1}) + \mathbf{G}_{k-1} \cdot (\mathbf{p}_{k-1} - \mathbf{p}_0)\right], \tag{10}$$

where $\mathbf{C_p}$ and $\mathbf{C_d}$ are the *a priori* covariance operators on parameters and data, respectively, $\mathbf{p}_0$ the starting model, and where $\mathbf{G}_{k-1} = \partial g(\mathbf{p}_{k-1})/\partial \mathbf{p}_{k-1}$ is the matrix of partial derivatives of the g function established in eqs (8) and (9). The indices related to $\mathbf{p}$ are now expressing the iteration rank and no longer the mode-branch radial order. Defining the kth image of the misfit function by

$$S(\mathbf{p}_k) = \frac{1}{2}\left\{[g(\mathbf{p}_k) - \mathbf{d}]^{\mathrm{T}} \cdot \mathbf{C_d}^{-1} \cdot [g(\mathbf{p}_k) - \mathbf{d}] + (\mathbf{p}_k - \mathbf{p}_0)^{\mathrm{T}} \cdot \mathbf{C_p}^{-1} \cdot (\mathbf{p}_k - \mathbf{p}_0)\right\}, \tag{11}$$

the maximum-likelihood point is defined by the minimum of $S(\mathbf{p})$. Minimizing the misfit function is then equivalent to finding the best compromise between decreasing the distance between the data vector and the image of the parameter vector through the g function, in the data space on one hand (first part of eq. 11), and not increasing the distance between the starting and the kth model on the other hand (second part of eq. 11), following the covariances defined in the *a priori* operators on the data and the parameters.

3.5.2 A priori data covariance operator

The *a priori* covariance operator on data, referred to as $\mathbf{C_d}$, includes data errors and also all effects that cannot be modelled by the g function defined in eq. (8) and (9). The only way to really measure each data error and then to compute realistic covariances in the data space, would be to obtain exactly the corresponding seismogram in which the signal due to the seismic event is removed. Hence, errors over the data space are impossible to determine correctly. In order to introduce as little *a priori* information as possible, the $\mathbf{C_d}$ matrix is computed with a constant value of 0.04 (including data and theory uncertainties) for the diagonal elements and zero for the off-diagonal elements. In other words, this choice means that the phase velocity perturbations are expected to explain at least 80 per cent of the recorded signal.

3.5.3 A priori parameter covariance operator

In the model space, the *a priori* covariance operator on parameters, referred to as $\mathbf{C_p}$, controls possible variations between the model vector components for a given iteration k (eq. 10), and also between the starting and the kth model (eq. 11). Considering that the phase velocity perturbation between two adjoining components (which are ordered according to increasing frequency values) of a given mode-branch do not vary too rapidly, $\mathbf{C_p}$ is a non-diagonal matrix. This *a priori* information reduces the number of independent components and then induces smoothed phase velocity perturbation curves. A typical behaviour of our problem resides in the way the parameter space is discretized. In the matrix domain, the distance between two adjoining components is always the same, whereas, as the model space is not evenly spaced

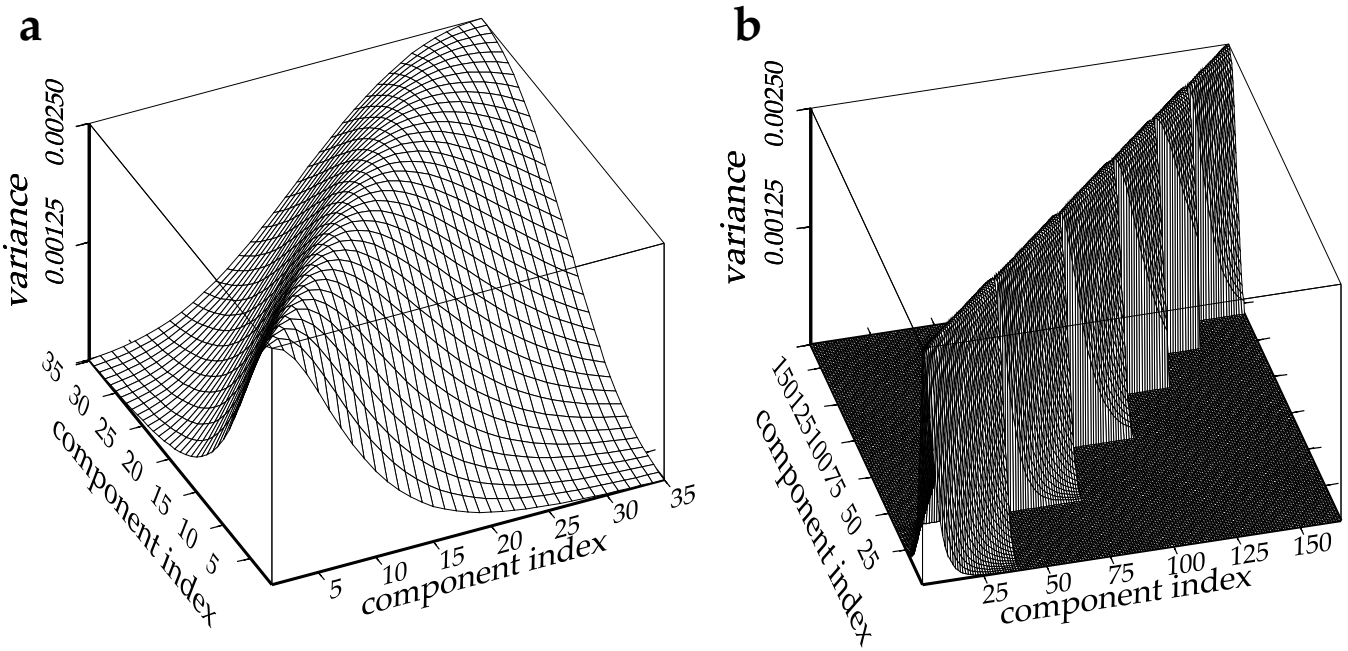

Figure 7. Illustration of two prior covariance operators on parameters. (a) The *x* and *y* axes represent the FM parameter vector component indices. As mentioned in Fig. 6(a) the first components correspond to the low-frequency values and the last ones to the high-frequency values. (b) In the HM case, the parameter vector is composed of the six selected frequency ranges ordered end to end, according to increasing frequency values inside each subvector. The number of parameter vector components is then equal to the number of mode-branches considered multiplied by the number of selected frequencies for each of them. The *x* and *y* axes represent the HM parameter vector component indices. In order not to impose mathematical relations between overtone phase velocities, correlations between two different mode-branch parameter subspaces are fixed to zero.

in the period sense, it differs when considering two frequencies next to each other. Thus, for a given mode-branch subvector, choosing a classical Gaussian distribution (defined by the component indices) as a covariance matrix over the model space leads to not imposing on the phase velocity perturbations the same physical *a priori* information over the whole frequency range. Consequently, the first components (low-frequency values) are strongly correlated, whereas the last ones (high-frequency values) are not correlated.

To avoid this feature, due to our choice to work in the spectral domain, and to impose the same physical *a priori* information over the whole model space, the correlation lengths are computed in terms of distance in the frequency (or period) sense. Thus, the *a priori* covariance operator on parameters is positive definite and expressed as

$$\mathbf{C}_{\mathbf{p}_{ij}} = M \exp\left[-\frac{1}{2}\frac{(i-j)^2}{\sigma_{ij}^2}\right], \tag{12}$$

where M is the maximum variance value authorized between two given components i and j. The only difference with respect to a classical Gaussian distribution resides in the expression of the correlation length σ_{ij} (or standard deviation). Let us denote by ν_i, the value of the ith frequency value and ξ as a frequency (or period) ratio. The standard deviation is defined as

$$\sigma_{ij} = \frac{(\nu_i + \nu_j)}{2\xi(\nu_{i+1} - \nu_i)}. \tag{13}$$

Practically, the M value is set to 0.0025 and the ξ frequency ratio to 0.5 (see Fig. 7a for the FM case). This means that, for each iteration, the phase velocity perturbations are expected within 5 per cent of the previous iteration value and for a given component, the correlation length is always equal to the half of the considered period value. This choice is consistent with the weak phase velocity perturbation assumption (see Section 2) and the large-scale exploration (see Section 3.4). These *a priori* constraints condition the inversion but do not limit the solution range, in the strict sense. If the data impose it, the model can get out of this expected domain but, consequently, the second part of the misfit value (eq. 11) is greatly increased, whereas the first part decreases. The role of such an approach is to keep as far as possible a good balance between the two terms defining the image of the misfit function (eq. 11). In the case of HM, in order to be consistent with the no *a priori* large-scale exploration of the model space, correlations between two different mode-branch components are fixed to zero, as shown in Fig. 7(b).

3.5.4 Multiple local optimizations

As summarized in Fig. 2, the vector corresponding to the best large-scale solution is used as a starting model to perform the first least-squares optimization. For example, in the FM case, it corresponds to the model that underlies the solution represented by the fourth circle in Fig. 6(b). Technically, the matrix inversion (eq. 10) is performed using a singular-value decomposition. Effectively, as the selection of

the significant frequency ranges is based on the synthetic amplitude spectra, it is possible that some parameter vector components are not correctly constrained. The real mode-branch amplitude spectra that do not necessarily match the synthetic and unresolved parts of the vector can then influence the well-constrained components. At each iteration k, the image of the corresponding model through the misfit function is determined. Optimizations stop when $|S(\mathbf{p}_k - S(\mathbf{p}_{k+1})| \leq \varepsilon$ is reached (with ε sufficiently small), and the last misfit value is stored. As we have already mentioned, a 2π phase skip can still remain over the whole frequency range of the measured phase velocity curves. For this reason, a new least-squares optimization is performed, this time starting from another previously selected model (for example, in the FM case, the model corresponding to the solution represented by the third circle in Fig. 6b). In the HM case, the large-scale model is composed of six subvectors and a difference of one of these is enough to induce several different phase velocity curves. At the end of each optimization, the corresponding misfit is stored, and a new inversion is performed for each of the starting models that have not yet been tested. For example, for the HM, the maximum number of local optimizations that is required in order to test all possible influences of one overtone on another, is 2^6, if the large-scale exploration detects two possible starting models for each of the six mode-branches.

Finally, the global best-fitting model corresponds to the solution that achieves the minimum of all local smallest misfit values previously determined. This detailed examination of all large-scale possibilities is a way to combine least-squares optimizations, leading by some *a priori* information and a non-linear exploration, in order not to limit the solution around the reference model. Such a technique produces a robust and well-fitting solution.

3.6 Estimation of the phase velocities reliability

A very important part of measuring higher-mode phase velocities is to assess their reliability. For each mode-branch, the significant frequency range is determined using the reference model (see Section 3.3), it is then essential to post-validate this frequency domain and also the phase velocity accuracies themselves. When dealing with the high non-linearity of the inverse problem for the whole model space, as performed by van Heijst & Woodhouse (1997), linear assessments are limited to computing final model errors. However, in our case, when choosing one of the vectors detected during the large-scale exploration as a starting model, the least-squares optimization conditions are supposed to be locally fulfilled. Thus, following the expression stated by Tarantola (1987), the *a posteriori* covariance operator is defined as

$$\mathbf{C_{pp}} = \mathbf{C_p} - \mathbf{C_p} \cdot \mathbf{G}_{k-1}^{\mathrm{T}} \cdot \mathbf{M}^{-1} \cdot \mathbf{G}_{k-1} \cdot \mathbf{C_p}, \quad \text{with} \quad \mathbf{M} = \left[\mathbf{C_d} + \mathbf{G}_{k-1} \cdot \mathbf{C_p} \cdot \mathbf{G}_{k-1}^{\mathrm{T}}\right], \tag{14}$$

where k denotes the last iteration. To perform the $\mathbf{M}$ matrix inversion, a classical algorithm is used, in order to really take into account all of the parameter vector components. The *a posteriori* uncertainties on each component are then given by the square root of the $\mathbf{C_{pp}}$ diagonal elements. Therefore, this reliability estimation combines, for each mode-branch, information concerning the relevance of the selected frequency range with the accuracy of the phase velocity measurements.

Another criterion of the *a posteriori* reliability estimate can be performed by shifting the synthetic seismograms following eq. (6), and then measuring, in the time domain, the waveform differences with the data. However, the phase velocities are determined over different frequency ranges depending on each mode-branch. So, the phase shift can only be achieved in these different frequency domains, which then can induce some artefacts in the time domain. Moreover, as several data are taken into account simultaneously, the shifted synthetic seismograms are rather 'mean' traces corresponding to the common ray path (see Section 3.2), rather than a series of seismograms that can

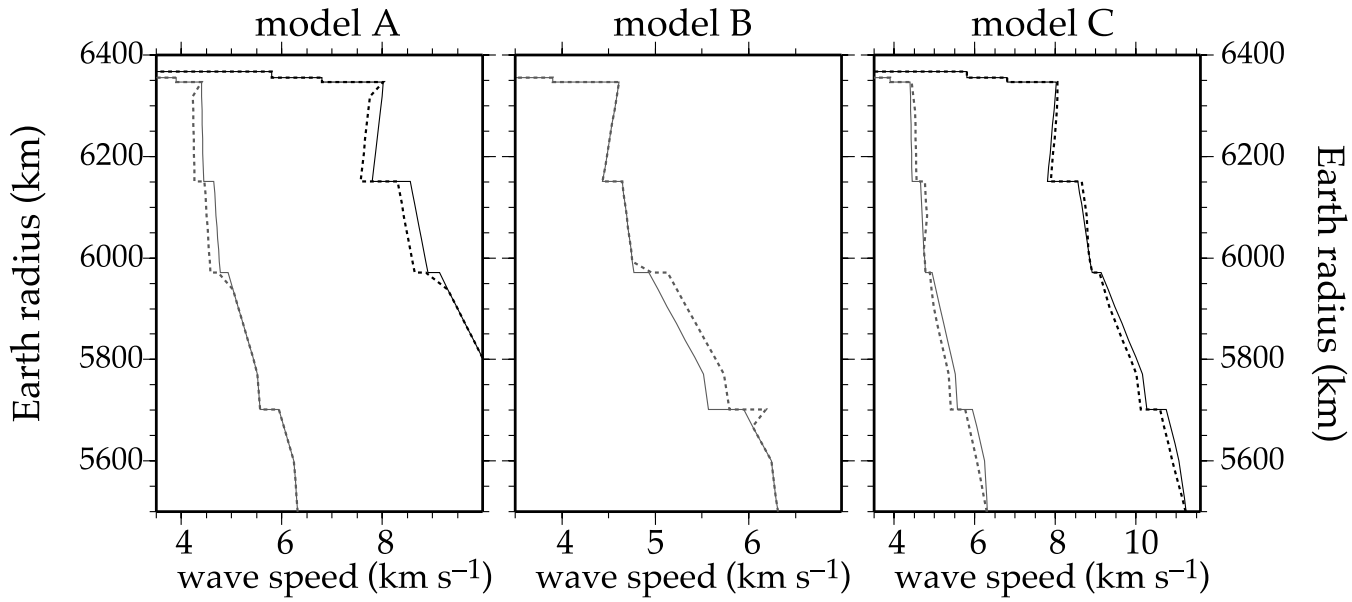

Figure 8. Different 1-D synthetic models, perturbed from the PREM, used for synthetic tests presented in Sections 4.2 and 4.3. Please refer to the Section 4.1 for details.

be directly compared one by one with each data point. For these reasons the reconstructed seismograms are qualitatively compared with the data, but the difference is not quantitatively measured. Only the variances given by eq. (14) are retained to *a posteriori* estimate the reliability of the measured phase velocities.

4 RESOLUTION TESTS USING SYNTHETIC DATA

In order to validate the roller coaster technique previously described, many synthetic tests have been performed. Only the results of four inversions using synthesized seismograms in atypical models are presented in this section.

4.1 Synthetic models

Three different models are used to synthesize data. They are displayed in Fig. 8 and detailed in the following. The *S*-wave velocity perturbations are applied both on S_V and on S_H waves.

(A) In model A, *P* and *S* PREM wave velocities are roughly reduced by 3 and 4 per cent, respectively, between 50 and 400 km depth.

(B) To simulate a deeper anomaly, the second model, hereafter referred to as model B, is identical to PREM except between 400 and 700 km depth. Within this range the *S* PREM wave velocity is roughly increased by 4 per cent.

(C) Model C is more complex. In the first anomaly, located between 80 and 300 km depth, *P* and *S* PREM wave velocities are smoothly increased by 1 and 2.5 per cent, respectively. Below this first anomaly, between 450 and 800 km depth, the *P* and *S* PREM wave velocities are smoothly reduced by 1.5 and 3 per cent, respectively.

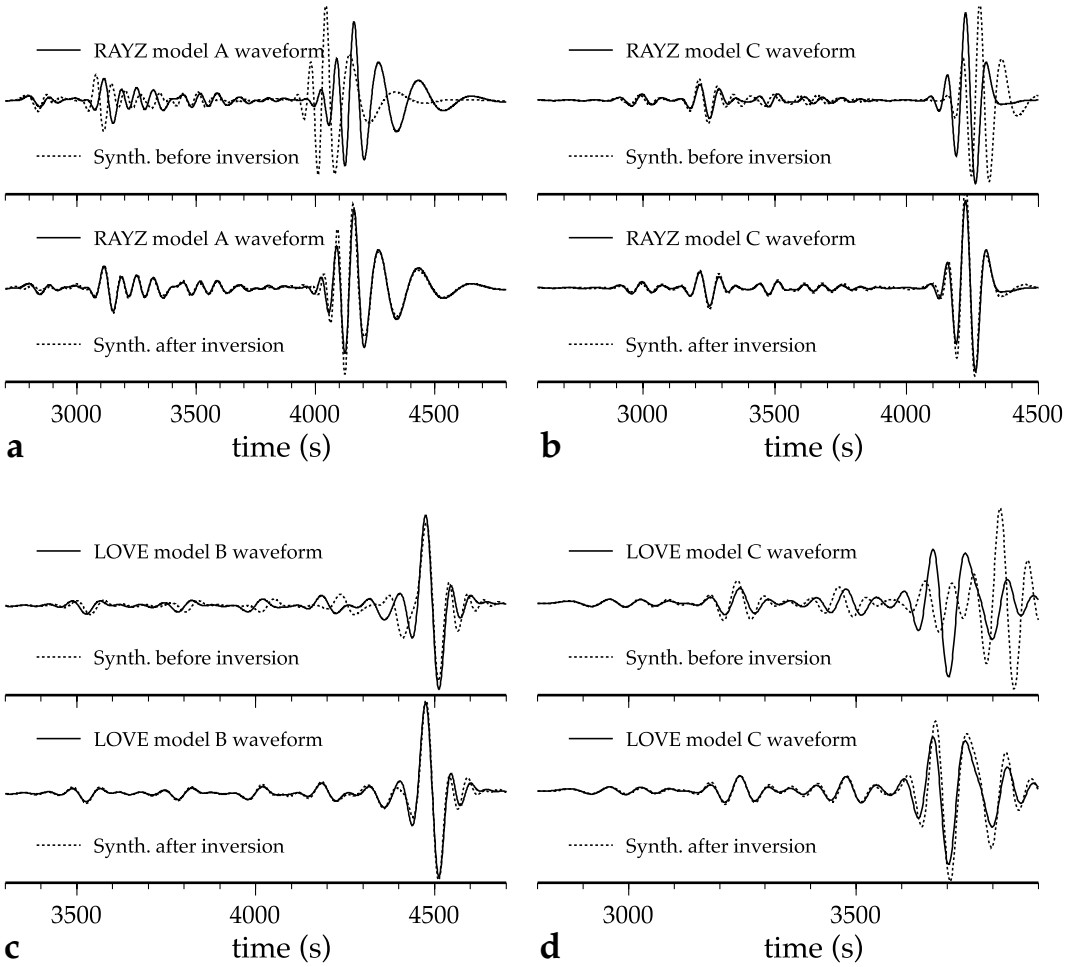

Figure 9. Synthetic test results corresponding to the phase velocity measurements presented in Fig. 10. For each test (see Sections 4.2 and 4.3 for details), one of the waveforms synthesized in a 1-D perturbed model (Fig. 8) is plotted as a solid line. The corresponding seismograms before (computed in PREM) and after inversion are drawn as dotted lines.

4.2 Rayleigh wave results

Concerning Rayleigh waves, two different tests are presented. For each of them, comparisons of one of the waveforms before and after inversion with the corresponding synthetic datum are shown in Figs 9(a) and (b). The results, in terms of phase velocity perturbation curves are displayed in Figs 10(a) and (b).

In the first one, six vertical traces are synthesized in model A (Fig. 8) and inverted simultaneously. Data correspond to six different events about 11 000 km far from the recording station and within a depth range of between 15 and 150 km. Real noise, recorded at the SCZ GEOSCOPE station (Stutzmann *et al.* 2000) is added. Although the model A heterogeneity is rather shallow, Fig. 9(a) shows that higher modes are greatly affected by such a velocity anomaly. As previously explained in Section 3.6 some artefacts, due to the different selected mode-branch frequency ranges and to the FM/HM time windowing, arise when shifting the PREM waveform using the measured phase velocity curves. This effect is visible in Fig. 9(a) around 4000 s. Nevertheless, phase velocities (Fig. 10a) are well recovered compared with the large discrepancy between the PREM phase velocities and those to retrieve, and also despite of the noise added to the synthetic data. Fig. 10(a)

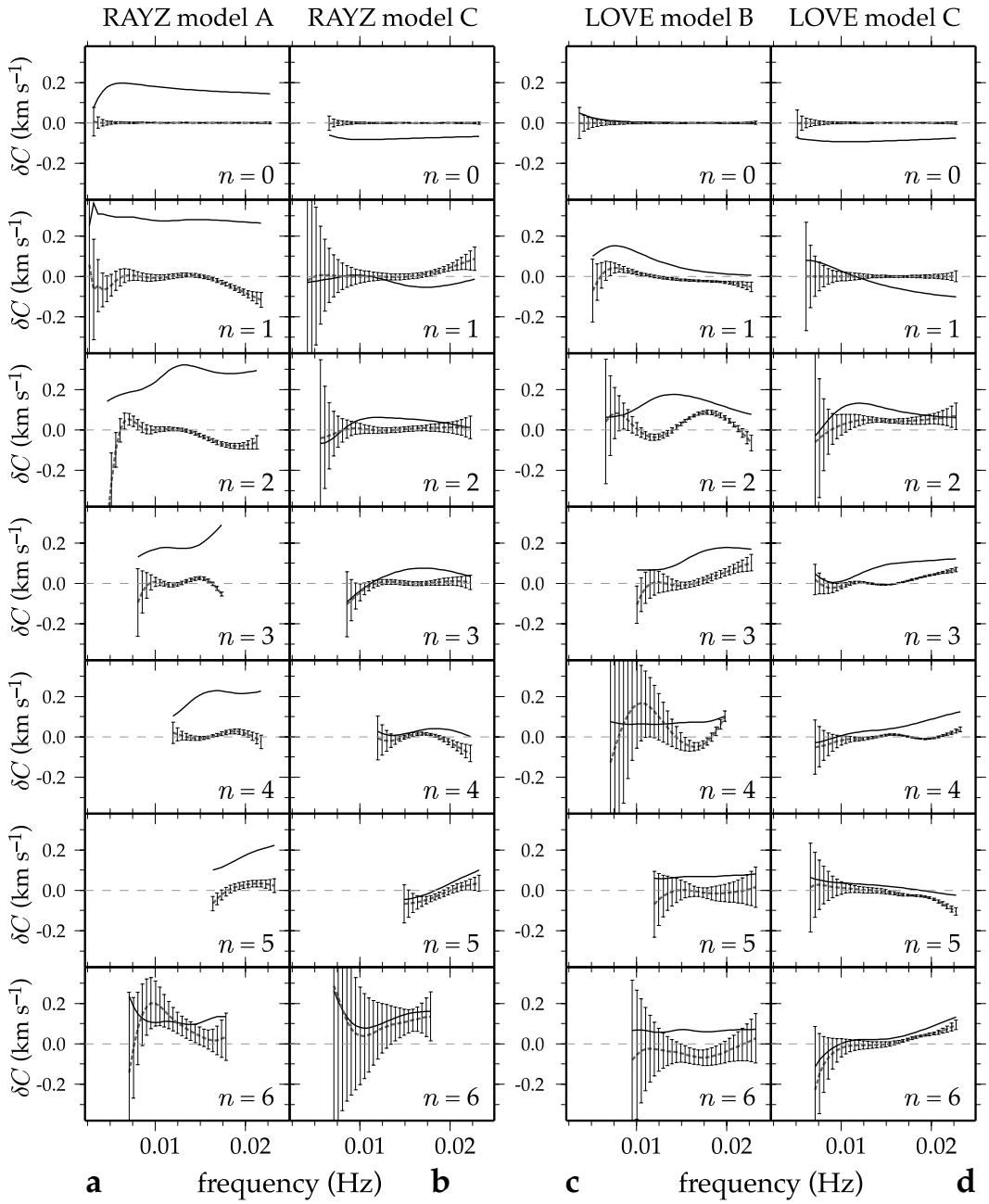

Figure 10. Synthetic test results, in terms of phase velocity perturbation curves, underlying the waveforms displayed in Fig. 9. For each mode-branch and each test (see Sections 4.2 and 4.3), the difference between the synthetic model (which has to be retrieved) and the reference phase velocities is drawn as a solid line. The difference between the synthetic model phase velocity and the inversion result is plotted as a grey dashed line with corresponding error bars.

Table 1. Numerical comparison of vertical Rayleigh wave results along the Vanuatu–California path displayed in Fig. 11. PREM velocity values are reported according to the period values defined by Cara (1979) (referred to as C.1979). These results are summarized in Lévêque & Cara (1985). The results obtained by van Heijst & Woodhouse (1997) (referred to as vHW.1997) are reported in the fifth column. Results of the present study are summarized in the last column.

Mode	Period (s)	PREM (km s^{-1})	C.1979 (km s^{-1})	vHW.1997 (km s^{-1})	This study (km s^{-1})
$n = 0$	110.70	4.127	4.102	4.139	4.135
	81.92	4.029	4.019	4.063	4.061
	57.69	3.967	3.986	4.036	4.033
$n = 1$	132.13	6.425	6.288	6.396	6.379
	95.26	5.782	5.673	5.745	5.710
	70.62	5.339	5.232	5.303	5.260
$n = 2$	53.90	5.815	5.694	5.771	5.750

shows that for a given mode-branch and for specific frequency values, phase velocity after inversion does not always match model A to the retrieved phase velocity. A major advantage of the roller coaster technique is that these mismeasurements are correctly taken into account in the *a posteriori* variances. It also appears that, focusing on the sixth overtone *a posteriori* uncertainties, this mode-branch is not sufficiently excited to accurately measure its phase velocity. This is further corroborated by an examination of its amplitude spectra (not displayed here).

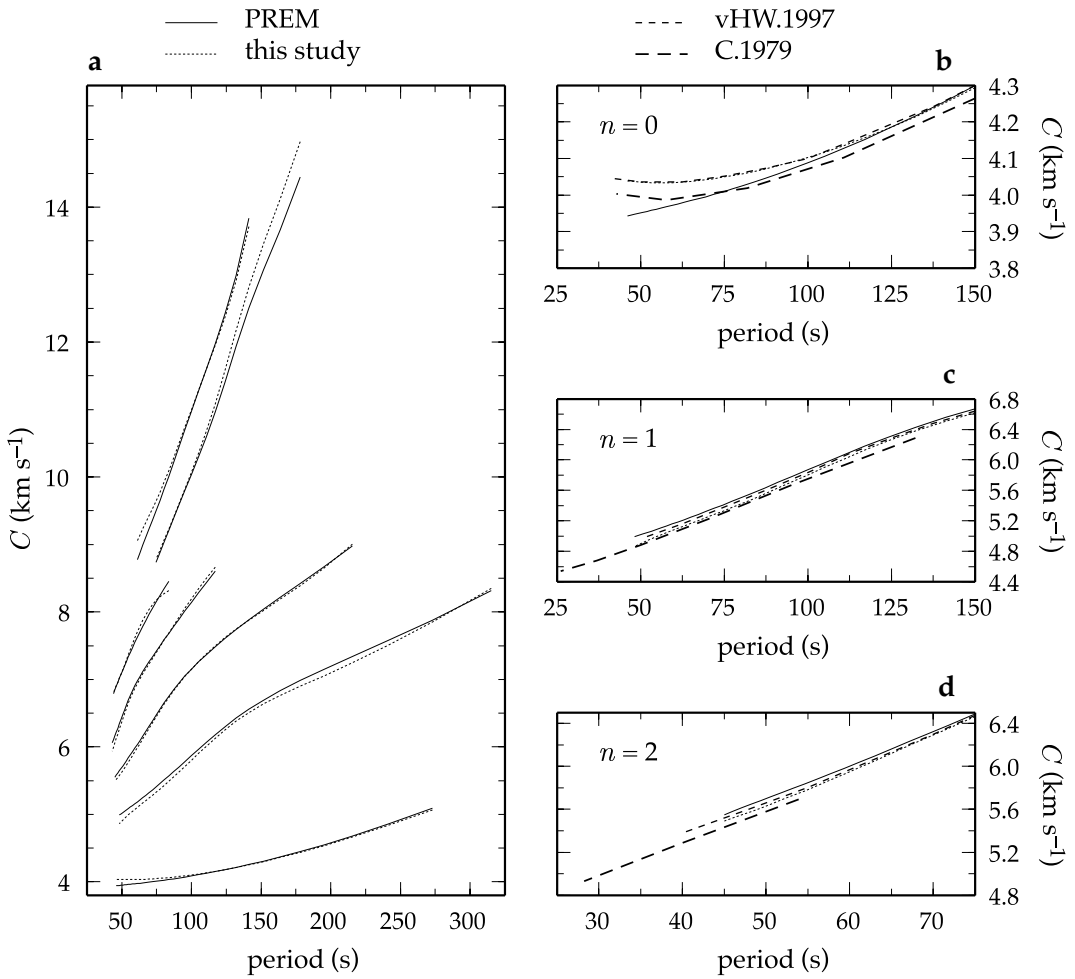

Figure 11. (a) Fundamental and higher-mode absolute phase velocities along the Vanuatu–California path. For each mode-branch, PREM phase velocities (solid lines) and measurements realized in this study (dotted lines) are plotted as a function of the period. The three seismograms inverted simultaneously in this study are displayed in Fig. 12. (b) Enlarged area of the left-side figure, for the fundamental mode. Our results are compared with those obtained by Cara (1979) (long-dashed line, referred to as C.1979) and by van Heijst & Woodhouse (1997) (short-dashed line, referred to as vHW.1997). (c) and (d) The same comparisons for the first and the second overtone, respectively.

The second test on Rayleigh waves consists in the inversion of four seismograms synthesized in model C (Fig. 8). Events underlying the waveforms are clustered about 11 550 km from the station and within a depth range of between 20 and 100 km. Results are displayed in Figs 9(b) and 10(b). The two superimposed anomalies induce some compensation effects on several HM phase velocities perturbations. Consequently, only the FM of the datum synthesized in model C (Fig. 9b), is largely different from the PREM synthetic seismogram. After inversion the whole signal waveform is correctly retrieved and due to the FM/HM time windowing (see Section 3.1), overtone weak amplitudes are not affected by the large FM amplitude. The phase velocities are correctly recovered (Fig. 10b) except for the sixth overtone, which is, however, taken into account well when considering its *a posteriori* variances.

4.3 Love wave results

Two different tests are presented concerning Love waves. The phase velocity perturbation curves are displayed in Figs 10(c) and (d). The comparisons between one of the synthetic seismograms, inversed simultaneously, with the initial and the final waveforms are shown in Figs 9(c) and (d).

In the first test, three different traces are synthesized in model B (Fig. 8). They correspond to three earthquakes, about 14 000 km away from the recording location and within a depth range between 15 and 35 km. The inversion result is displayed in Figs 9(c) and 10(c). Considering the model B roughness, the phase velocities are rather correctly recovered. The weak amplitude of the HM part of the signal implies that different overtones are inverted in the same too wide frequency range. Thus, the underdetermination of the problem is not sufficiently reduced and several phase velocities are not well constrained (see, for example, the fourth overtone). However, even if a given mode-branch phase velocity is not correctly measured, first the use of a SVD-type algorithm allows one to assess that such a mode-branch does not bias the determination of the others, and second the *a posteriori* estimation of the reliability reflects this unresolved overtone correctly.

The second test concerning Love waves consists in the inversion of seven seismograms synthesized in model C (Fig. 8). Earthquakes underlying the waveforms are located about 11 450 km from the station and within a depth range between 15 and 256 km. Considering the initial discrepancy between model C and the PREM waveform (Fig. 9d), the signal is correctly retrieved. As previously mentioned, the different overtones are inverted over a too wide frequency range (Fig. 10d). Thus, as a consequence of this underdetermination, several mode-branch phase velocities are not well measured (see, for example, the second overtone). However, once again, the use of an SVD inverse algorithm allows other mode-branches to not be affected by this mismeasurement and the corresponding *a posteriori* variances clearly point out unresolved overtones, which not underlie the waveforms in a significant way.

5 REAL DATA INVERSION RESULTS

In this section we present Rayleigh and Love phase velocity measurements along the Vanuatu Islands-to-California path.

The Rayleigh wave results are compared with those obtained by Cara (1979) and van Heijst & Woodhouse (1997), hereafter referred to as C.1979 and vHW.1997, respectively. The different results are displayed in Fig. 11 and part of them is summarized in Table 1. The seven different phase velocities (FM+HM) are measured by inverting three seismograms provided by earthquakes clustered in the Vanuatu region (13°S, 167°E) and recorded at the SCZ GEOSCOPE station. Their depth varies between 98 and 210 km. All the waveforms inverted simultaneously are shown in Fig. 12. In Fig. 11(a), FM and the first six overtone phase velocities are plotted as a function of period. For each of the three traces displayed in Fig. 12, the signal is well recovered and phase velocities are determined with a high reliability estimation (except for the fifth overtone low frequencies). For the FM (Fig. 11b), the phase velocity measured in this study is almost identical to that measured by vHW.1997. At short periods, between 45 and 100 s, the two curves differ noticeably from the PREM phase velocity. This upward

Table 2. Numerical comparison of Love wave results along the Vanuatu–California path. PREM velocity values are reported for several selected periods according to the periods expressed by Lévêque & Cara (1985) (denoted as LC.1985 in this table). The results of the present study are summarized in the last column.

Mode	Period (s)	PREM (km s^{-1})	LC.1985 (km s^{-1})	This study (km s^{-1})
$n = 0$	150.00	4.776	4.939	4.803
	110.00	4.668	4.780	4.707
	80.00	4.588	4.703	4.645
	60.00	4.528	4.658	4.611
$n = 1$	110.00	6.013	6.180	5.979
	80.00	5.503	5.608	5.494
	50.00	5.043	5.118	5.017
$n = 2$	80.00	6.667	6.830	6.598
	55.00	5.817	5.917	5.690

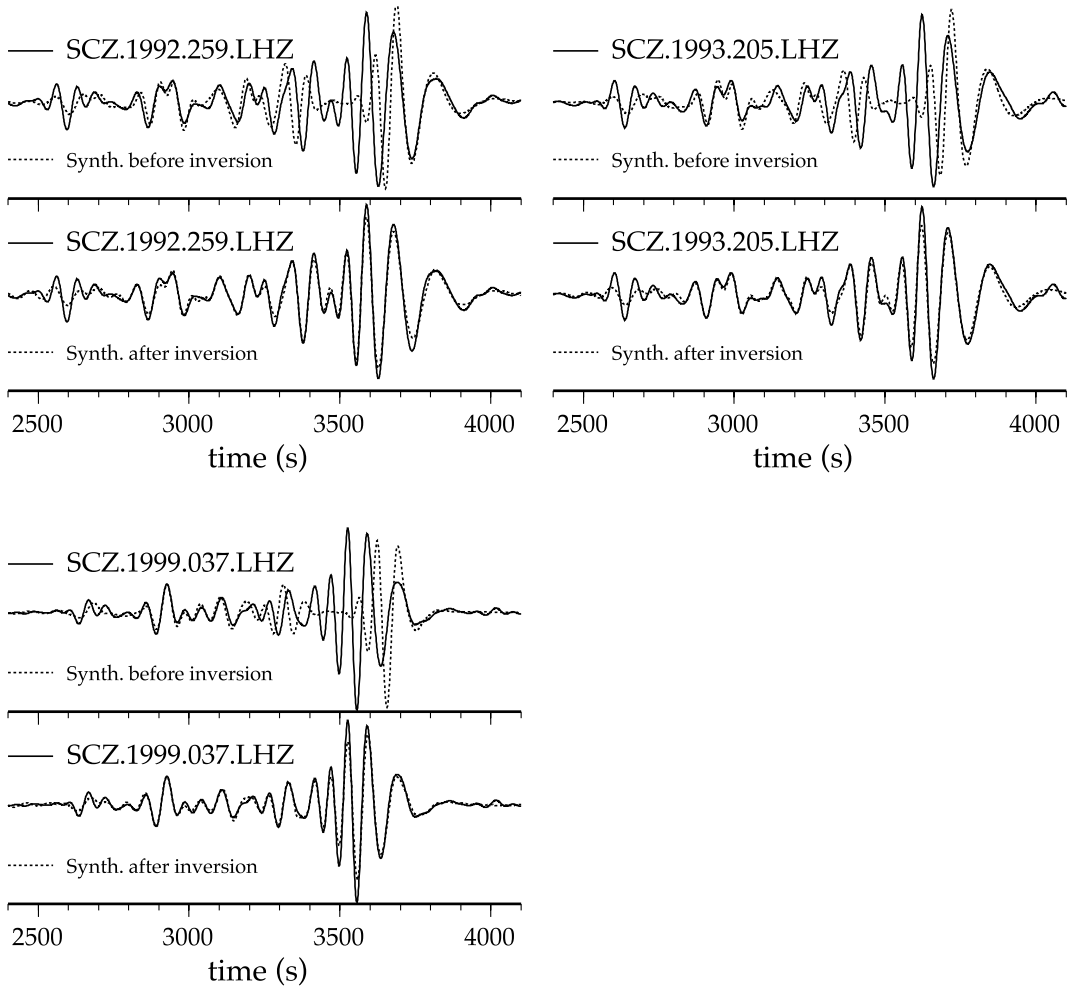

Figure 12. Waveforms used to measure Rayleigh FM and HM phase velocities along the Vanuatu–California path (Fig. 11). Each of the three seismograms (solid line) is compared with the synthetic waveforms before inversion (computed in PREM) and after phase shift using inversion results. Both are represented as a dashed line.

trend has already been noticed by C.1979 but to a lesser extent. Concerning the first overtone (Fig. 11c), our results agree correctly with those found by C.1979 and vHW.1997. It confirms the presence of a phase velocity negative perturbation at short periods. Between 50 and 100 s, our measurement is between the phase velocities measured by C.1979 and by vHW.1997. At longer periods, our result is closer to that of vHW.1997. For the second overtone (Fig. 11d), our and vHW.1997 results are quite similar. All the phase velocities are smaller than that from PREM, but when comparing with the C.1979 results, the size of the perturbations is different. Our and vHW.1997 perturbations are smaller than those of C.1979. However, the frequency ranges in which C.1979 measured the HM phase velocity differs significantly from the two other studies, which makes any comparison difficult.

Part of the Love wave results, for several selected period values, is summarized in Table 2. Our results (FM and the first two mode-branches) are compared with the phase velocities mentioned by Lévêque & Cara (1985). In this study, three seismograms are inverted simultaneously. Events are clustered in the Vanuatu region centred around 15°S, 167°E, which approximately reproduces the experimental conditions stated by Lévêque & Cara (1985). Then, the epicentral distance of the barycentre is approximately 9470 km. The different earthquake depths vary between 15 and 122 km. In contrast to Rayleigh wave phase velocity measurements, there is a large discrepancy between the results of Lévêque & Cara (1985) and ours. This contrast is probably due to stronger coupling effects between the different mode-branches. Concerning FM, this discrepancy tends to reduce with increasing frequency values but globally, at long periods, our results are closer to the PREM phase velocities than Lévêque & Cara (1985). A more detailed study over a larger frequency domain is then required in order to allow any warranted comparison.

6 CONCLUSION

The use of the roller coaster technique, presented in this paper, enables one to measure several mode-branch phase velocities efficiently, with a real non-linear algorithm. Each mode-branch solution, defined over its own significant frequency range, is a smoothed average phase

velocity along an epicentre–station path. These measurements explain well both real and synthetic Rayleigh and Love waveforms, within a maximum period range included between 40 and 500 s. Except for the mode-branch significant frequency range determination, the roller coaster technique is not sensitive to the reference model, for which synthetic seismograms are computed. In order to take this possible bias into account, a SVD algorithm is used for the matrices inversion. In this article, we do not improve the forward problem but highlight how it is important not to *a priori* constrict the solution around the reference model, which is even made when starting from it to perform a gradient least-squares inversion in order to solve a highly non-linear problem. The main characteristic of our approach is based on the simple idea of first determining several possible configurations for each mode-branch phase velocity, assuming large-scale variations over the frequency range, and secondly to test each of these configurations to finally retain the solution that achieves the minimum misfit. Such an algorithm is an alternative to combining some *a priori* information over the parameter space and an inversion scheme that preserves a large range of solutions. This method, since it combines a systematic large–scale exploration with multiple local least–squares optimizations, accounts for the high non–linearity and the nonunicity of the forward problem and is therefore greatly sensitive to the data set. In order to check the data quality, the procedure is not completely automated but the processing time is sufficiently reasonable to treat a large number of seismograms rapidly. In the future, this approach can be used with a more realistic wave perturbation theory, including effects such as cross-branch coupling and/or off-great-circle propagation. In every synthetic test we present here, each mode-branch *a posteriori* reliability estimate turns out to be a very powerful instrument in assessing the phase velocity measurements. Concerning Rayleigh waves, our Vanuatu–California path results and previously determined ones seem to correctly agree. A more detailed study and comparisons with other measurements, using different methods, are required for Love waves.

ACKNOWLEDGMENTS

We would like to thank the operators of the GEOSCOPE network for providing high-quality waveform data. We are very grateful to Barbara Romanowicz and Éric Debayle for their constructive reviews and interesting suggestions. The authors sincerely acknowledge Michel Cara for making the 1979 numerical values available to us. EB would like to thank Alex Fournier for useful corrections and for finding the name 'rollercoaster', and Geneviève 'LaTeX wizard' Moguilny. All figures were produced using the GMT graphics software.

REFERENCES

Cara, M., 1979. Lateral variations of *s*-velocity in the upper mantle from higher Rayleigh modes, *Geophys. J. R. astr. Soc.,* **57,** 649–670.

Cara, M. & Lévêque, J.J., 1987. Waveform inversion using secondary observables, *Geophys. Res. Lett.,* **14,** 1046–1049.

Debayle, É., 1999. s_v-wave azimuthal anisotropy in the Australian upper mantle: preliminary results from automated Rayleigh waveform inversion, *Geophys. J. Int.,* **137,** 747–754.

Deuss, A. & Woodhouse, J.H., 2001. Theoretical free-oscillation spectra: the importance of wide band coupling, *Geophys. J. Int.,* **146,** 833–842.

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

Ekström, G. & Dziewonski, A.M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature,* **394,** 168–172.

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

Giardini, D., Li, X. & Woodhouse, J.H., 1987. Three-dimensional structure of the Earth from splitting in free oscillation spectra, *Nature,* **325,** 405–411.

Grand, S.P., van der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth, *GSA Today,* **7,** 1–7.

Kanamori, H. & Given, J.W., 1981. Use of long-period surface waves for rapid determination of earthquake-source parameters, *Phys. Earth planet. Inter.,* **27,** 8–31.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.,* **101,** 16 059–19 075.

Laske, G. & Masters, G., 1998. Surface-wave polarization data and global anisotropic structure, *Geophys. J. Int.,* **132,** 508–520.

Lebedev, S., 2000. The upper mantle beneath the western Pacific and southeast Asia, *PhD thesis,* Princeton University, Princeton.

Lebedev, S. & Nolet, G., 2003. Upper mantle beneath southeast asia from *s* velocity tomography, *J. geophys. Res.,* **108,** 2048, doi:10.1029/2000JB000073.

Lévêque, J.J. & Cara, M., 1985. Inversion of multimode surface wave data: evidence for sub-lithospheric anisotropy, *Geophys. J. R. astr. Soc.,* **83,** 753–773.

Lévêque, J.J., Cara, M. & Rouland, D., 1991. Waveform inversion of surface wave data: test of a new tool for systematic investigation of upper mantle structures, *Geophys. J. Int.,* **104,** 565–581.

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

Li, X.D. & Tanimoto, T., 1993. *Waveform Inversion of Long-Period Seismic Data for Global Structure,* pp. 64–91, Chapman & Hall, London.

Liu, X.F. & Dziewonski, A.M., 1998. *Global Analysis of Shear Wave Velocity Anomalies in the Lower-Most Mantle,* pp. 21–36, American Geophysical Union, Washington.

Lognonné, P. & Romanowicz, B., 1990. Modelling of coupled normal modes of the earth: the spectral method, *Geophys. J. Int.,* **102,** 365–395.

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,** 709–728.

Montagner, J.P., 1986. Regional three-dimensional structures using long-period surface waves, *Ann. Geophys.,* **4,** 283–294.

Montagner, J.P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.,* **95,** 4797–4819.

Nataf, H.C., Nakanishi, I. & Anderson, D.L., 1986. Measurements of mantle velocities and inversion for lateral heterogeneities and anisotropy III, *J. geophys. Res.,* **91,** 7261–7307.

Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.,* **95,** 8499–8512.

Nolet, G., van Trier, J. & Huisman, R., 1986. A formalism for non linear inversion of seismic surface waves, *Geophys. Res. Lett.,* **13,** 26–29.

Okal, E., 1978. A physical classification of the Earth's spheroidal modes, *J. Phys. Earth,* **26,** 75–103.

Ritsema, J. & van Heijst, H.J., 2000. Seismic imaging of structural heterogeneity in Earth's mantle: evidence for large-scale mantle flow, *Sci. Progr.,* **83,** 243–259.

Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 1999. Complex shear wave velocity structure imaged beneath africa and iceland, *Science,* **286,** 1925–1928.

Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle, *J. geophys. Res.,* **100,** 12 375–12 394.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm—II. appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Stutzmann, É. & Montagner, J.P., 1993. An inverse technique for retrieving higher mode phase velocity and mantle structure, *Geophys. J. Int.,* **113,** 669–683.

Stutzmann, É. & Montagner, J.P., 1994. Tomography of the transition zone from the inversion of higher-mode surface waves, *Phys. Earth planet. Inter.,* **86,** 99–115.

Stutzmann, É., Roult, G. & Astiz, L., 2000. Geoscope station noise levels, *Bull. seism. Soc. Am.,* **90,** 690–701.

Su, W.J., Woodward, R.L. & Dziewonski, A.M., 1994. Degree 12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.,* **99,** 6945–6980.

Tanimoto, T., 1990. Long wavelength *s*-velocity structure throughout the mantle, *Geophys. J. Int.,* **100,** 327–336.

Tarantola, A., 1987. *Inverse Problem Theory,* Elsevier, Amsterdam.

Tarantola, A. & Valette, B., 1982a. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.,* **20,** 219–232.

Tarantola, A. & Valette, B., 1982b. Inverse problems = quest for information, *J. Geophys.,* **50,** 159–170.

Trampert, J. & Woodhouse, J.H., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.,* **122,** 675–690.

van der Hilst, R.D., Widyantoro, S. & Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography, *Nature,* **386,** 578–584.

van Heijst, H.J. & Woodhouse, J.H., 1997. Measuring surface-wave overtone phase velocities using a mode-branch stripping technique, *Geophys. J. Int.,* **131,** 209–230.

van Heijst, H.J. & Woodhouse, J.H., 1999. Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode-branch stripping, *Geophys. J. Int.,* **137,** 601–620.

van Heijst, H.J., Ritsema, J. & Woodhouse, J.H., 1999. Global *p* and *s* velocity structure derived from normal mode splitting, surface wave dispersion and body wave travel time data, in *EOS, Trans. Am. geophys. Un.,* Vol. 80, AGU Fall Meet. Suppl.

Woodhouse, J.H., 1980. The coupling and attenuation of nearly resonant multiplets in the Earth's free oscillation spectrum, *Geophys. J. R. astr. Soc.,* **61,** 261–283.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89,** 5953–5986.

Woodhouse, J.H. & Dziewonski, A.M., 1986. Three-dimensional mantle models based on mantle wave and long period body wave data, *EOS, Trans. Am. geophys. Un.,* **67,** 307.

Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized Earth model, *Geophys. J. R. astr. Soc.,* **68,** 653–673.

Woodhouse, J.H. & Wong, Y.K., 1986. Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. astr. Soc.,* **87,** 753–773.

Yoshizawa, K. & Kennett, B.L.N., 2002. Non-linear waveform inversion for surface waves with a neighbourhood algorithm—application to multimode dispersion measurements, *Geophys. J. Int.,* **149,** 118–133.

Zhang, Y. & Tanimoto, T., 1991. Global Love wave phase velocity variation and its significance to plate tectonics, *Phys. Earth planet. Inter.,* **66,** 160–202.

APPENDIX A: THIRD-ORDER DEVELOPMENT

In this appendix, we show that the forward problem expressed in eq. (6) is equivalent to an expression using a phase velocity perturbation but expanded to third order. Considering the *j*th mode-branch, we assume that

$$\delta C_j(\mathbf{r},\omega) = C_j^{(\mathrm{R})}(\mathbf{r},\omega) - C_j^{(\mathrm{S})}(\omega), \tag{A1}$$

then eq. (2) rises as

$$A^{(\mathrm{R})}(\mathbf{r},\omega)\exp\left[i\Phi^{(\mathrm{R})}(\mathbf{r},\omega)\right] = \sum_{j=0}^{n} A_j^{(\mathrm{S})}(\mathbf{r},\omega)\exp\left\{i\left[\Psi_j(\mathbf{r},\omega) - \frac{\omega a\Delta}{C_j^{(\mathrm{S})}(\omega)}\chi'\right]\right\}, \tag{A2}$$

with

$$\chi' = \left[1 + \frac{\delta C_j(\mathbf{r},\omega)}{C_j^{(\mathrm{S})}(\omega)}\right]^{-1}, \tag{A3}$$

$\mathbf{r}$, ω, A, Φ, a, Ψ and Δ as previously defined in Section 2. Introducing a new expression for the dimensionless parameter vector, denoted as $\mathbf{p'}_j(\mathbf{r},\omega)$ and defined as follows:

$$\mathbf{p}'_j(\mathbf{r},\omega) = \frac{C_j^{(\mathrm{R})}(\mathbf{r},\omega) - C_j^{(\mathrm{S})}(\omega)}{C_j^{(\mathrm{S})}(\omega),} \tag{A4}$$

expanding the χ' term up to third order and introducing the synthetic phase $\Phi_j^{(\mathrm{S})}(\mathbf{r},\omega)$, as the sum of the source phase and the phase shift due to the propagation in the reference model, the forward problem becomes

$$A^{(\mathrm{R})}(\mathbf{r},\omega)\exp\left[i\Phi^{(\mathrm{R})}(\mathbf{r},\omega)\right] = \sum_{j=0}^{n} A_j^{(\mathrm{S})}(\mathbf{r},\omega)\exp\left[i\left\{\Phi_j^{(\mathrm{S})}(\mathbf{r},\omega)\frac{\Delta}{C_j} + \frac{\omega a\Delta}{C_j^{(\mathrm{S})}(\mathbf{r},\omega)}\left[\mathbf{p}'_j(\mathbf{r},\omega) - \mathbf{p}'^2_j(\mathbf{r},\omega) + \mathbf{p}'^3_j(\mathbf{r},\omega)\right]\right\}\right]. \tag{A5}$$

If the choice to develop the χ' term (eq. A3) to first order is made, then the forward problem can be written as

$$A^{(\mathrm{R})}(\mathbf{r},\omega)\exp\left[i\Phi^{(\mathrm{R})}(\mathbf{r},\omega)\right] = \sum_{j=0}^{n} A_j^{(\mathrm{S})}(\mathbf{r},\omega)\exp\left\{i\left[\Phi_j^{(\mathrm{S})}(\mathbf{r},\omega) + \frac{\omega a\Delta}{C_j^{(\mathrm{S})}(\omega)}\mathbf{p}'_j(\mathbf{r},\omega)\right]\right\}. \tag{A6}$$

To test the accuracy of eqs (A5) and (A6) a very simple experiment is performed and is only presented for a single mode-branch for reasons of clarity (Fig. A1). A fundamental mode waveform is synthesized in model A (see Section 4.1 and Fig. 8 for more details), assessing that phase

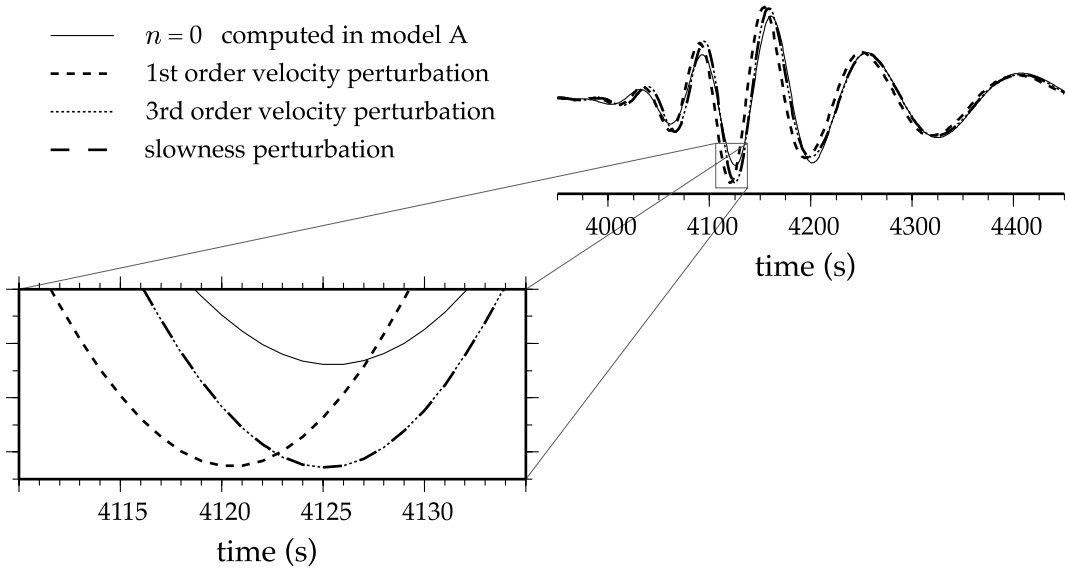

Figure A1. The dotted line represents a fundamental mode seismogram computed in model A (see Section 4.1 and Fig. 8 for more details). The model A phase velocity is used directly to phase shift the FM seismogram computed in PREM (not displayed here for clarity). When expanding the phase velocity perturbation to first order (eq. A6) the reconstructed signal (small-dashed line) is phase shifted by 5 s as shown in the zoomed part of the figure. It is correctly retrieved when expanding the phase velocity perturbation up to third order (eq. A5) (represented as a dotted line). The long-dashed line represents the recovered phase when using a slowness perturbation development (eq. 6). Details are presented in Appendix A.

velocity perturbations are lower than 5 per cent. The direct computation of eigenfunctions in both models (PREM and model A) enables one to obtain the exact expression of the parameter $\mathbf{p}'$ (eq. A4), over the whole frequency range (between 40 and 500 s). Thus, starting from the PREM synthetic seismogram and shifting the phase term as shown in eqs (A5) and (A6), provides two waveforms that should match the model A seismogram phase exactly. In Fig. A1 it clearly appears that a first-order development induces a 5 s phase shift, whereas the waveform is correctly recovered when developing the χ' term up to third order.

The same test is performed in order to compare the waveforms shifted following eqs (6) and (A5) and the model A seismogram. The result is displayed in Fig. A1. The phase of the model A signal is correctly retrieved for both cases. The largest difference between third-order velocity perturbation development (eq. A5) and slowness perturbation development (eq. 6) waveforms is lower than 0.001 per cent. Several other tests have been performed with larger perturbation size (until 10 per cent) and reinforce this result. In the perturbation range used in this study, eqs (6) and (A5) are then strictly equivalent.

APPENDIX B: INFLUENCE OF DIFFERENT TIME WINDOW DIMENSIONS

In order to test the influence of the time picking step in the roller coaster technique, results of several experiments are presented in this appendix. Three different seismograms, clustered in the same area, are inverted (see Section 5). They are located near the Vanuatu islands within a depth range between 98 and 210 km, and are recorded at the SCZ GEOSCOPE station. For a given trace, five different time pickings are considered. The influence of these pickings is summarized in the Table B1, for one of these seismograms. The different picking labels are explained in Section 3.1. In the first four columns, the variation of the pickings are reported with respect to the initial picking time. We present in the last three columns, θ as defined hereafter, for three different mode-branches (fundamental mode, first and third overtone):

$$\theta_n = \frac{1}{l_{\max}} \sum_{l=1}^{l_{\max}} C^1_{l,n} - C^i_{l,n},$$

where n denotes the mode-branch considered, $l_{\max}$ is the number of frequencies and where i is the test number. These experiments show that the phase velocity estimates are very robust for a wide range of time window dimensions. This is a consequence of the simultaneous use of several seismograms, and the using of an SVD algorithm that stabilizes the inversion results.

Table B1. Five different time pickings realized for a given real seismogram and the corresponding average of differences over the inverted frequency domain, for three mode-branches. For each line the results are presented with respect to the initial test. See Appendix B for more details.

Test no	δT1 (s)	δT2 (s)	δT3 (s)	δT4 (s)	θ_0 (km s^{-1})	θ_1 (km s^{-1})	θ_3 (km s^{-1})
2	501.7	−108.9	−78.1	−147.5	1.48×10^{-4}	4.37×10^{-4}	-5.24×10^{-3}
3	500.7	−109.1	−131.4	−143.6	-1.17×10^{-3}	2.99×10^{-3}	-3.98×10^{-3}
4	522.0	−52.4	−41.4	66.0	5.26×10^{-4}	2.34×10^{-3}	-1.02×10^{-3}
5	0.0	−108.9	−131.4	0.0	5.13×10^{-5}	2.25×10^{-3}	-1.62×10^{-3}