Comparison of global waveform inversions with and without considering cross-branch modal coupling

Xiang-Dong Li and Barbara Romanowicz
Seismographic Station, University of California, Berkeley, CA 94720, USA

SUMMARY

In this study we present a new approach to inverting long-period seismic waveforms. The effect of lateral heterogeneity is partitioned into two. The first part represents the effect of horizontally averaged structure along the great circle between the source and receiver, and is allowed to remain in non-linear form in the formulation. The second part incorporates any further correction due to cross-branch modal coupling, which has been neglected in the more conventional path average approximation (PAVA). This term is linearized and then treated asymptotically so that the seismogram depends only upon the structure within the great-circle section determined by the source and receiver (Li & Tanimoto 1993a). We refer to this new method as the *non-linear asymptotic coupling theory* (NACT). The sensitivity kernels predicted by the PAVA and NACT are compared. While the sensitivity kernels are similar for surface waves and shallow-turning body waves, they are very different for body waves that sample the deep mantle. By examining the inversion algorithms for the PAVA and NACT, we demonstrate that the computation time required by the NACT tends to be of the same order of magnitude as that required by the PAVA, as the number of model parameters increases. Based upon a realistically large data set (5041 body-wave seismograms and 1531 mantle-wave seismograms), formal resolution analyses are performed using both PAVA and NACT. We find that the NACT is significantly more powerful in resolving 3-D structure in the deep mantle. We compare the models obtained for the same observed data set using the two approaches. As expected, they differ more in the lower mantle than in the upper mantle. The difference in their amplitude spectra increases with spherical harmonic degree. The model developed using the NACT predicts the observed surface geoid better than that developed using the PAVA, based upon geodynamic flow modelling.

Key words: Earth's interior, inversion, mode coupling, seismic tomography, synthetic waveforms.

1 INTRODUCTION

Recent tomographic models of the Earth's elastic structure rely primarily on either traveltime data of first-arriving phases collected by the International Seismological Centre (e.g. Inoue *et al.* 1990; Pulliam, Vasco & Johnson 1993) or waveforms of long-period surface and body waves (e.g. Tanimoto 1990; Woodward & Masters 1991a,b; Masters, Bolton & Shearer 1992; Zhang & Tanimoto 1993; Dziewonski *et al.* 1993; Su, Woodward & Dziewonski 1994). The advantages of using digital data of long-period waves, now that many broad-band stations are operating around the world, comes from the high quality of measurements, essentially free of the noise introduced by unchecked reading errors, and the possibility of better sampling the volume of the mantle through the inclusion of many later arriving phases. In addition, the error due to unmodelled small-scale structure should be smaller for long-period waves because of their longer wavelengths. The techniques used in many waveform studies are based on the so-called path average approximation (PAVA), under which the seismogram is sensitive only to the horizontally averaged structure along the great circle between the source and receiver. Such an approximation works fairly well for

surface waves, but is inaccurate for body waves, because it fails to describe the concentration of sensitivity to structure in the vicinity of the ray path. An alternative approach has been developed (Woodward & Masters 1991a,b; Su & Dziewonski 1991), in which differential or absolute traveltimes of long-period body-wave phases are measured from waveforms and then interpreted in the framework of geometrical ray theory. Although the simplicity of this approach is very attractive in many ways, its limitations are obvious. Traveltimes are difficult to measure for overlapping phases, which are common for long-period waves and sample some important regions not reachable by well-isolated phases. In addition, the geometrical ray theory, which is accurate in the high-frequency limit, may be inadequate when applied to long-wavelength seismic waves. In the hope that the PAVA waveform fitting and the ray approach would be complementary to each other, Dziewonski & Woodward (1992), Woodward *et al.* (1993), and Su *et al.* (1994) combined the two methods and inverted waveforms and traveltime measurements of long-period seismic waves simultaneously. A more recent study (Liu, Su & Dziewonski 1994) demonstrates that the addition of *SKS-S* traveltime measurements can further improve the resolution at the bottom of the mantle.

It is obvious that waveform inversions based upon more accurate theories are desirable. In principle, the exact solution to the equation of motion in a laterally heterogeneous earth model can be approached using the variational method or its equivalents. Significant progress in developing more efficient algorithms using these exact theories has been made (Hara, Tsuboi & Geller 1991, 1993; Geller & Hara 1993). At the same time, higher order perturbation theories have also been developed (e.g. Dahlen 1987; Lognonné & Romanowicz 1990; Park 1990). Nevertheless, computations required by these approaches are still heavy for very large-scale inversions using current computational facilities. It is, therefore, useful to continue investigating asymptotic approaches.

In this study we present a new approximate technique for waveform inversion, based upon a theoretical work by Li & Tanimoto (1993a). Taking cross-branch coupling into account, the theory is able to bring out the ray character of body waves using normal-mode superposition. It represents a more accurate approach than the PAVA, while keeping the amount of computations relatively small.

The original version of the modal couping theory presented in Li & Tanimoto (1993a) involves the complete linearization of phase perturbation terms, an approximation which may not be valid beyond very short times. In this study, we introduce a method that overcomes this deficiency, at the cost of introducing non-linear terms in the theory and having to solve the inverse problem iteratively. We shall refer to this *non-linear asymptotic coupling theory* as NACT in this paper.

We apply the NACT to the inversion of a data set consisting of approximately 5000 long-period *SH* body-wave and 1500 long-period *SH* mantle-wave seismograms. The inversion result is compared with that obtained using the more conventional PAVA base upon the same data set. It has been demonstrated (Mochizuki 1986a; Park 1987; Romanowicz 1987) that the PAVA is equivalent to neglecting cross-branch modal coupling. Therefore the

comparison of the PAVA and NACT shows how significant cross-branch coupling is in global waveform inversions. We shall restrict the scope of this paper to the comparison of the two inversion techniques and defer the discussion of geodynamical implications of our model to a separate publication.

2 NON-LINEAR ASYMPTOTIC COUPLING THEORY OF NORMAL MODES

Based upon first-order perturbation theory for normal modes (e.g. Woodhouse 1983), Li & Tanimoto (1993a) developed an asymptotic formulation to calculate seismograms of long-period seismic waves. Eq. (18) of Li & Tanimoto (1993a), which was given in an incorrect index notation, should be written in a matrix notation,

$$u(\tau) = \mathbf{R} \exp (i\Omega\tau)\mathbf{S}, \tag{1}$$

where $u(\tau)$ is a seismogram as a function of time τ, the real part of the right side of the equation is understood, $\mathbf{R}$ and $\mathbf{S}$ are the receiver and source vectors, respectively, as defined in Woodhouse & Girnius (1982) with some convention differences (see Li & Tanimoto 1993a, for details), and the Ω can be defined through its elements

$$\Omega_{ij} = \omega_i I_{ij} + \frac{Z_{ij}}{\omega_i + \omega_j}. \tag{2}$$

In eq. (2) ω_i is the (complex) eigenfrequency of singlet i of the reference spherically symmetrical earth model (which degenerates to the same value for all singlets belonging to the same multiplet), I_{ij} are the elements of the identity matrix $\mathbf{I}$, Z_{ij} represent the effect of 3-D structure perturbation as defined in Li & Tanimoto (1993a). The matrix exponentiation needs to be evaluated through a power series

$$\exp \mathbf{A} = \mathbf{I} + \mathbf{A} + \frac{1}{2}\mathbf{A}^2 + \cdots + \frac{1}{n!}\mathbf{A}^n + \cdots \tag{3}$$

In Li & Tanimoto (1993a) the non-linear dependence of seismogram $u(\tau)$ on the structure effect Z_{ij} was directly linearized under the assumption that $Z_{ij}\tau$ are small, which is inaccurate especially for large τ (Um, Dahlen & Park 1991). Here we present a method to improve the accuracy. Let us introduce apparent frequency shifts $\delta\omega_i$ which are the same for all singlets belonging to the same multiplet and, generally speaking, are functionals of the source–receiver distribution and of the given 3-D earth model. We can then rewrite eq. (2):

$$\Omega_{ij} = (\omega_i + \delta\omega_i)I_{ij} + \frac{Z_{ij}}{\omega_i + \omega_j} - \delta\omega_i I_{ij} \tag{4}$$

$$\equiv \hat{\omega}_i I_{ij} + \frac{\hat{Z}_{ij}}{\omega_i + \omega_j}, \tag{5}$$

where

$$\hat{\omega}_i \equiv \omega_i + \delta\omega_i, \tag{6}$$

and

$$\hat{Z}_{ij} \equiv Z_{ij} - 2\omega_i\delta\omega_i I_{ij}. \tag{7}$$

Following Woodhouse (1983), we may define $\delta\omega_k$, which is

the same for all the singlets belonging to the same multiplet, as

$$\delta\omega_k = \frac{\sum\limits_{m,m'} R_m^k Z_{mm'}^k S_{m'}^k}{2\omega_k \sum\limits_m R_m^k S_m^k}, \tag{8}$$

where the absolute singlet indices i,j used above, which label all the singlets globally, have been switched to a multiplet index k and relative singlet indices m,m', which represent different singlets within a multiplet (thus singlet indices can be dropped for quantities which do not vary with singlets within a multiplet). However, we here elect to calculate $\delta\omega_k$ using the path average method,

$$\delta\omega_k = \frac{1}{\widehat{SR}} \int_S^R \delta\omega_{local}^k \, d\varphi, \tag{9}$$

where the integration is taken along the great-circle path from the source S to receiver R, $d\varphi$ denotes the incremental angular distance element, $\widehat{SR}$ is the total length of the path, and $\delta\omega_{local}^k$ is the local frequency introduced in Jordan (1978). The main advantage of eq. (9) is its numerical simplicity. The physical meaning of partitioning in eq. (5) is also very clear when eq. (9) is used: the first term on the right side of eq. (5) implies that, to the first order, a seismogram in a laterally heterogeneous earth model can be calculated in the same way as in a spherically symmetric model using a 1-D theory, providing that the eigenfrequencies $\hat{\omega}_k$ are evaluated for a model obtained by horizontal average along the great circle connecting the source and receiver; the second term represents any further correction, whose effect is presumably much smaller and can thus be linearized. This results in a better approximation than complete linearization of the total effect. Replacing ω_k by $\hat{\omega}_k$ and Z_{ij} by $\hat{Z}_{ij}$, we can go through the same procedure as in Li & Tanimoto (1993a) and obtain, after asymptotic approximation and some algebra,

$$u(\tau) = \sum_k A_k \exp(i\hat{\omega}_k \tau) + u_1(\tau), \tag{10}$$

where the summation is taken over all the multiplets, $A_k = \sum_m R_m^k S_m^k$, and

$$u_1(\tau) = -\sum_k i\tau\delta\omega_k A_k \exp(i\hat{\omega}_k \tau)$$

$$+ \sum_k \sum_{k' \in \Gamma_k} D_{kk'}(\tau) E_{kk'}, \tag{11}$$

with Γ_k being the set of multiplets whose eigenfrequencies $\omega_{k'}$ are higher than or equal to ω_k. In eq. (11) $D_{kk'}$ and $E_{kk'}$ are defined by

$$D_{kk'}(\tau) = \frac{\exp(i\hat{\omega}_k \tau) - \exp(i\hat{\omega}_{k'}\tau)}{(\omega_k + \omega_{k'})(\hat{\omega}_k - \hat{\omega}_{k'})}, \tag{12}$$

where the limit $\hat{\omega}_k \rightarrow \hat{\omega}_{k'}$ is understood for $\hat{\omega}_k = \hat{\omega}_{k'}$, and

$$E_{kk'} = \frac{1}{2\pi}\left[Q_{kk'}^{(1)} \int_0^{2\pi} \delta\omega_{kk'}^2 \cos j\varphi \, d\varphi \right.$$

$$\left. + Q_{kk'}^{(2)} \int_0^{2\pi} \delta\omega_{kk'}^2 \sin j\varphi \, d\varphi \right], \tag{13}$$

where $Q_{kk'}^{(1)}$ and $Q_{kk'}^{(2)}$ are functionals of the source and receiver and are given in Appendix A, $j \equiv l' - l$ is the difference between the angular degrees l' and l (of multiplet k' and k, respectively), and φ is the distance along the great circle. $\delta\omega_{kk'}^2$ in eq. (13) may be calculated through

$$\delta\omega_{kk'}^2 = \int_0^a \delta\mathbf{m} \cdot \mathbf{M}_{kk'}(r) r^2 \, dr - \sum_d r_d^2 h_d H_{kk'}^d, \tag{14}$$

where a is the Earth's radius, $\delta\mathbf{m}$ represents the volumetric perturbations in earth structure (the vector notation represents possible multiple physical quantities, such as P and S velocities and density), and h_d is the undulation of the dth discontinuity. If anisotropy of earth structure is ignored, the kernels $\mathbf{M}_{kk'}$ and $H_{kk'}^d$ may be evaluated directly using eqs (A36)–(A42) of Woodhouse (1980) (where the case $l'' = 0$ is required). When anisotropy is considered the expressions for $\mathbf{M}_{kk'}$ and $H_{kk'}^d$ can be modified in a straightforward manner (Mochizuki 1986b; Tanimoto 1986b; Romanowicz & Snieder 1988; Li 1990).

In summary, the non-linear asymptotic coupling theory (NACT), expressed in eq. (10), states that a seismogram can be calculated in two parts: the first term on the right-hand side of eq. (1) is the exact result of the PAVA (see Tanimoto 1986a; for equivalence between the formulations of Tanimoto 1986a and Woodhouse & Dziewonski 1984, see Li & Tanimoto 1993b); the second term represents the cross-branch modal coupling effect ignored by the PAVA.

3 COMPARISON OF SENSITIVITY KERNELS PREDICTED BY THE PAVA AND NACT

A natural question arises as to how significant the difference between the PAVA and NACT is. In this section we attempt to answer the question by comparing the sensitivity kernels predicted by the two theories. Following Li & Tanimoto (1993a) we define the sensitivity kernel $\mathbf{M}$ of a seismogram through

$$\Delta u(\tau) = \int_S \mathbf{M}(r, \varphi; \tau)\delta\mathbf{m}(r, \varphi) \, dA, \tag{15}$$

where Δu is the perturbation seismogram contributed from a volumetric perturbation $\delta\mathbf{m}$, which is sampled by the sensitivity kernel $\mathbf{M}$, the surface integration is over the whole area S of the great-circle section $(dA = r \, d\varphi \, dr)$. A sensitivity kernel predicts the regions that a particular seismic phase, arriving at the receiver at time τ, has sampled along its path.

As an example, a synthetic SH seismogram at epicentral distance $\Delta = 110°$ of a hypothetical earthquake with a source depth of 35 km is calculated using a spherically symmetric earth model (PREM of Dziewonski & Anderson 1981) and is plotted in Fig. 1, where phases S_{diff}, SS and G are also indicated. The seismogram is calculated using the modal superposition technique (Gilbert & Dziewonski 1975), which is equivalent to the first term on the right-hand side of eq. (10) with $\hat{\omega}_k = \omega_k$. Our aim here is to compare the
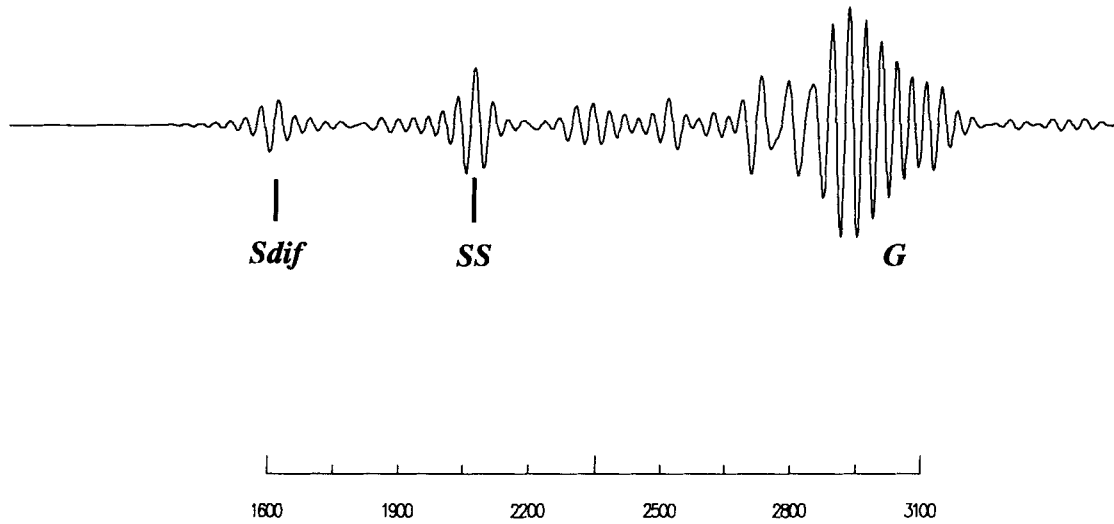
Figure 1. An *SH* synthetic seismogram at epicentral distance $\Delta = 110°$. The seismogram is calculated for a spherically symmetric reference model (PREM of Dziewonski & Anderson 1981) and low-pass filtered with a cut-off frequency $f = 1/32$ Hz and a corner frequency $f = 1/38$ Hz. The S_{dif}, *SS*, and *G* phases are indicated. The time-scale is in seconds.

sensitivity of the seismograms to structure perturbation as predicted by the two theories. For *SH* seismograms, we assume that they are sensitive only to *S*-velocity perturbation, and the sensitivity kernel **M** thus reduces to a scalar function M. In Fig. 2 are shown the sensitivity kernels predicted by the PAVA and NACT for *G* (top), *SS* (middle), and S_{dif} (bottom) phases. The sensitivity kernels in the left-hand column are calculated using the PAVA and those in the right-hand column are calculated using the NACT. From visual inspection of these kernels, the two theories predict similar sensitivity kernels for surface waves but very different ones for body waves. If we measure the difference between the kernel M_P predicted by the PAVA and the kernel M_N calculated using the NACT by the quantity

$$V = \frac{\int_S (M_P - M_N)^2 \, dA}{\sqrt{\int_S M_P^2 \, dA \int_S M_N^2 \, dA}},$$ (16)

where the integrations are over the area of the great-circle section, we obtain $V = 6$, 103 and 197 per cent for *G*, *SS*, S_{dif}, respectively. The deviation of the two approaches becomes more severe as the wave penetrates deeper into the mantle.

On the other hand, the difference between the NACT and geometrical ray theory is also clear from Fig. 2. First, the NACT predicts finite-width Fresnel zones associated with body-wave propagation. Secondly, the kernels from the NACT indicate a non-uniform sensitivity distribution along the ray path. The sensitivity concentrates at the source and receiver regions and at the bouncing point for *SS* or diffraction region for S_{dif}. Such concentrations at the source and receiver regions are also shown in independent studies (e.g. Stark & Nikolayev 1993). In particular, the sensitivity concentration near the Earth's surface has the important

consequence that the crustal structure has a stronger effect on body-wave seismograms than predicted by the PAVA approach and the geometrical ray theory.

4   INVERSION ALGORITHMS

Since the coupling between multiplets is taken into account in the NACT formulation (see the double summation in the second term on the right-hand side of eq. 11), the computation time seems, at first glance, to be forbiddingly expensive for large-dimension inverse problems. However, we shall show in this section that if spherical harmonics are used to expand the earth model, coupling will be limited only to those multiplets whose angular degrees satisfy a certain selection rule. We will further demonstrate that the computation times required by the two approaches, the NACT and PAVA, are asymptotically the same, as the maximum spherical harmonic degree of lateral heterogeneity increases.

We note that in both the PAVA and NACT, a seismogram depends only upon the structure within the great-circle section containing the source and receiver. For a given combination of the source and receiver, we may define 'great-circle coordinates' (ϑ, φ) in such a way that the source is located at $\vartheta = \pi/2$ and $\varphi = 0$ and the receiver is located at $\vartheta = \pi/2$ and $\varphi = \Delta$ with $\Delta < \pi$ (asymptotic theories like the PAVA and NACT break down at $\Delta = 0$ and $\Delta = \pi$, where caustics of travelling waves occur). Within the great circle (the equator of the great-circle coordinates), we may express a 3-D model $\delta\mathbf{m}$ in terms of Fourier series, $\{\cos h\varphi, \sin h\varphi\}$, and certain radial basis functions $\mathbf{f}_q(r)$ (again, the vector notation represents possible multiple physical quantities),

$$\delta\mathbf{m}|_{\vartheta=\pi/2} = \sum_q \sum_h (a_h^q \cos h\varphi + b_h^q \sin h\varphi)\mathbf{f}_q(r),$$ (17)
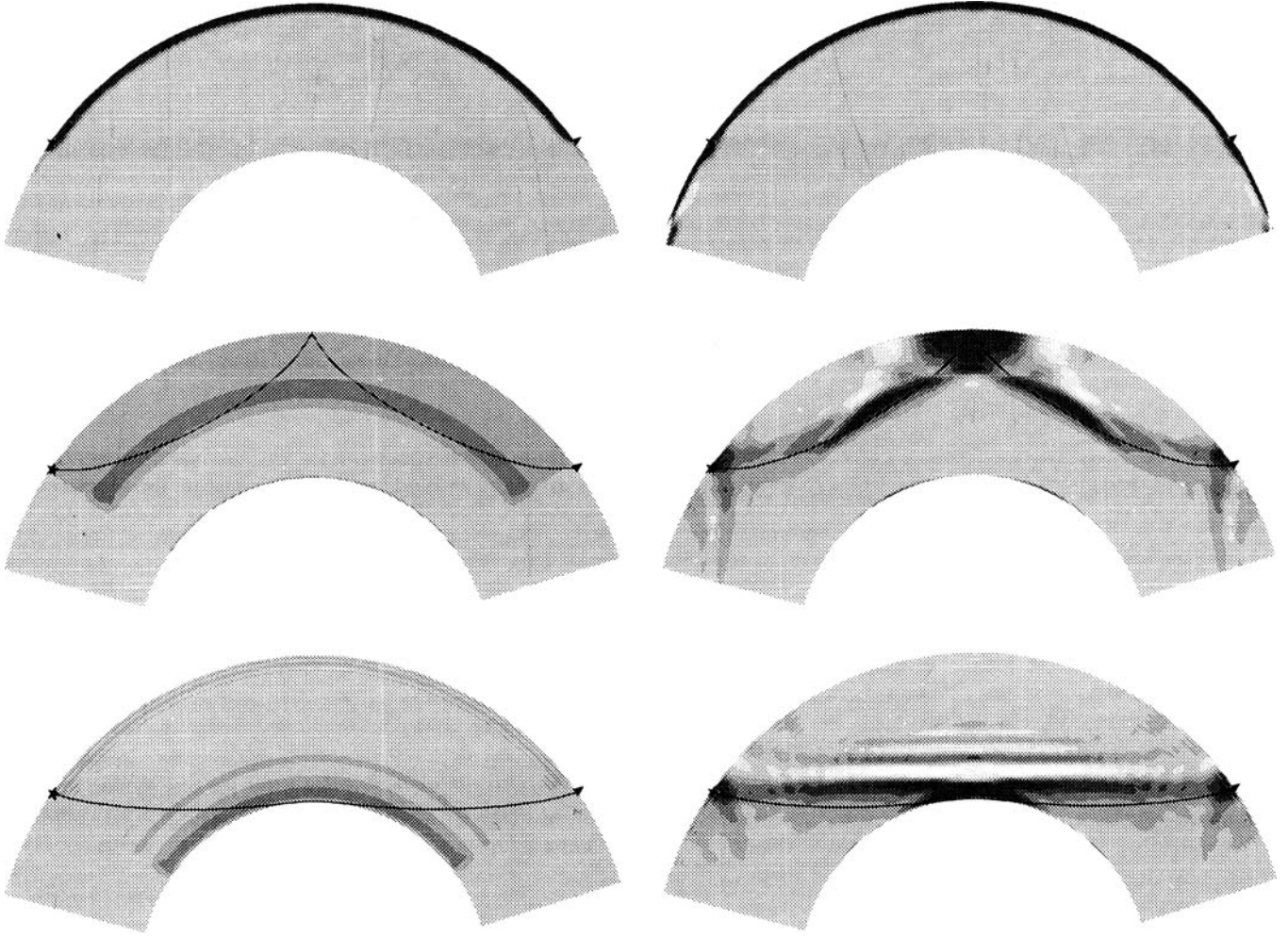
Figure 2. Sensitivity kernels of the seismogram shown in Fig. 1. The earthquake hypocentre is represented by stars and the station is represented by triangles. The geometrical rays shown as thin curves are calculated using the geometrical ray theory. The kernels are for the G (top), SS (middle), and S_{dif} (bottom) phases. The kernels shown on the left are calculated using the PAVA and the ones shown on the right are calculated using the NACT. The graphic contour intervals are 4:2:1 from the top down.

where coefficients a_h^q and b_h^q characterize the model. Using eqs (10)–(14) and (17), we can express a seismogram in terms of model coefficients a_h^q and b_h^q,

$$u(\tau) = \sum_k \left\{ A_k \exp(i\hat{\omega}_k \tau)(1 - i\tau\delta\omega_k) \right.$$
$$\left. + \sum_q \sum_h \left(a_h^q \sum_{k' \in \Gamma_{kh}} A_{kk'}^q + b_h^q \sum_{k' \in \Gamma_{kh}} B_{kk'}^q \right) \right\}, \qquad (18)$$

where Γ_{kh} denotes the set of multiplets k' that have reference eigenfrequency $\omega_{k'} \geq \omega_k$ and angular degree l' satisfying $|l' - l| = h$, and coefficients $A_{kk'}^q$ and $B_{kk'}^q$ are given in Appendix B. Now we can see that the summation over k' in eq. (11) is taken over all the multiplets with reference frequencies $\omega_{k'} \geq \omega_k$, while in eq. (18) the summation is only over those that satisfy the additional selection rules $|l' - l| = h$, resulting from orthogonality of sine and cosine functions.

An explicit relationship between the apparent frequency

shift $\delta\omega_k$ and coefficients a_h^q and b_h^q can be obtained from eqs (9), (14) and (17),

$$2\omega_k\delta\omega_k = \sum_q M_{kk}^q \left[a_0^q - (-1)^n \right.$$
$$\left. \times \frac{1}{\widehat{SR}} \sum_{h>0} \left(\frac{\sin h\Delta}{h} a_h^q + \frac{1-\cos h\Delta}{h} b_h^q \right) \right], \qquad (19)$$

where n is the orbit number of the wave packet (for common minor-arc body-wave phases $n = 1$), M_{kk}^q are given in Appendix B, and $\widehat{SR}$ is given by

$$\widehat{SR} = \begin{cases} (n-1)\pi + \Delta, & \text{if } n = 1, 3, 5, \ldots \\ n\pi - \Delta, & \text{if } n = 2, 4, 6, \ldots \end{cases} \qquad (20)$$

Thus eqs (18) and (19) provide an algorithm for calculating a synthetic seismogram from model parameters a_h^q and b_h^q. The partial derivatives of a seismogram with respect to

parameters a_h^q and b_h^q can be readily derived:

$$\frac{\partial u}{\partial a_h^q} = \sum_k \left\{ i\tau A_k \exp(i\hat{\omega}_k \tau) \frac{\partial \delta \omega_k}{\partial a_h^q} + \sum_{k' \in \Gamma_{kh}} A_{kk'}^q \right.$$

$$+ \left[(\tau \delta \omega_k - i)\tau A_k \exp(i\hat{\omega}_k \tau) \frac{\partial \delta \omega_k}{\partial a_h^q} \right.$$

$$\left. \left. + \sum_{h'} \sum_{k' \in \Gamma_{kh'}} \left(F_{kk'}^{h'k} \frac{\partial \delta \omega_k}{\partial a_h^q} + F_{kk'}^{h'k'} \frac{\partial \delta \omega_{k'}}{\partial a_h^q} \right) \right] \right\}, \qquad (21)$$

and

$$\frac{\partial u}{\partial b_h^q} = \sum_k \left\{ i\tau A_k \exp(i\hat{\omega}_k \tau) \frac{\partial \delta \omega_k}{\partial b_h^q} + \sum_{k' \in \Gamma_{kh}} B_{kk'}^q \right.$$

$$+ \left[(\tau \delta \omega_k - i)\tau A_k \exp(i\hat{\omega}_k \tau) \frac{\partial \delta \omega_k}{\partial b_h^q} \right.$$

$$\left. \left. + \sum_{h'} \sum_{k' \in \Gamma_{kh'}} \left(F_{kk'}^{h'k} \frac{\partial \delta \omega_k}{\partial b_h^q} + F_{kk'}^{h'k'} \frac{\partial \delta \omega_{k'}}{\partial b_h^q} \right) \right] \right\}, \qquad (22)$$

where $\partial \delta \omega_k / \partial a_h^q$ and $\partial \delta \omega_k / \partial b_h^q$ are obtainable from eq. (19) and

$$F_{kk'}^{h'k''} = \frac{\partial D_{kk'}}{\partial \delta \omega_{k''}} \sum_p (U_{kk'} a_{h'}^p + V_{kk'} b_{h'}^p) M_{kk'}^p \qquad (23)$$

with $U_{kk'}$ and $V_{kk'}$ being given in Appendix B. We note that in eq. (21), and similarly in eq. (22), the first term corresponds to the PAVA; the second term alone is the result of the linear asymptotic coupling theory of Li & Tanimoto (1993a).

Although a single seismogram is sensitive to, and hence constrains, only the structure within the great-circle section, we can model 3-D structure using many seismograms, which sample different great-circle sections. Let us expand a 3-D earth model $\delta \mathbf{m}$ in terms of spherical harmonics $Y_s^t(\theta, \phi)$ (Edmonds 1960), where θ and ϕ are the geographic coordinates, and the radial basis functions $\mathbf{f}_q(r)$:

$$\delta \mathbf{m} = \sum_q \sum_{s=0}^{s_{max}} \sum_{t=-s}^{s} c_{stq} Y_s^t(\theta, \phi) \mathbf{f}_q(r), \qquad (24)$$

where c_{stq} are coefficients to be determined, we truncate the spherical harmonic expansion at a finite angular degree s_{max}, and the dimension of the radial basis functions is also finite. It can be demonstrated that the 'great-circle parameters' a_h^q and b_h^q are related to the 'global parameters' c_{stq} through

$$a_h^q = \sum_{s=h}^{s_{max}} \sum_{t=-s}^{s} \alpha_{st}^h c_{stq}, \qquad (25)$$

$$\underset{s+h \text{ even}}{}$$

and

$$b_h^q = \sum_{s=h}^{s_{max}} \sum_{t=-s}^{s} \beta_{st}^h c_{stq}, \qquad (26)$$

$$\underset{s+h \text{ even}}{}$$

where α_{st}^h and β_{st}^h are, for a given source–receiver distribution, constant and can be readily calculated using the transformation property of spherical harmonics under the rotation of the coordinate frame (Edmonds 1960). Thus eqs (25) and (26), together with eqs (18) and (19), give a recipe for calculating seismograms from c_{stq}. For inverse problems, the partial derivatives of a seismogram u with respect to c_{stq}

can be obtained using the chain rule of differentiation:

$$\frac{\partial u}{\partial c_{stq}} = \sum_{\substack{h=0 \\ s+h \text{ even}}}^{s} \left(\alpha_{st}^h \frac{\partial u}{\partial a_h^q} + \beta_{st}^h \frac{\partial u}{\partial b_h^q} \right). \qquad (27)$$

Now let us examine the number of operations involved in inversions using the PAVA and NACT, as a function of s_{max}. Both theories are non-linear and need to be solved iteratively. Within each iteration, computations are mainly used in three steps:

(1) calculating the synthetic seismograms and partial derivatives with respect to a_h^q and b_h^q;

(2) constructing the partial derivatives with respect to c_{stq}; and

(3) solving the linearized least-squares problem.

The difference between the PAVA and NACT exists only in the first step. In the first step, the computations required by the PAVA depend upon s_{max} only through the fact that the total number of partial derivatives a_h^q and b_h^q increases with s_{max}. Since $\partial u / \partial a_h^q$ and $\partial u / \partial b_h^q$ need to be calculated for all $h = 0, 1, \ldots, s_{max}$ (see eq. 27), the computation time required by the PAVA in the first step is proportional to s_{max}. For the NACT, the computations for each term of $\partial u / \partial a_h^q$ and $\partial u / \partial b_h^q$ are proportional to s_{max}, because of the summations over h' in eqs (21) and (22). Thus the total computations required by the NACT in the first step are proportional to s_{max}^2. The second step involves the evaluation of eq. (27). For a given s, the computation time is proportional to s^2 since $-s \le t \le s$ and $h \le s$. The total computation time is, therefore, proportional to $(\sum_{s=0}^{s_{max}} s^2) \to s_{max}^3$. The computations required in the third step also increase faster than s_{max}^2, as in the first step. For example, if we solve the linearized least-squares problem by using the normal equations, the computations needed in constructing the inner product matrix (often known as $\mathbf{A}^T\mathbf{A}$ matrix in the literature) increase as s_{max}^4. For the inversion experiments performed in this study, where s_{max} is set to 8, approximately 70 per cent of the total computation time used in the NACT approach is consumed in the first step due to the modal superposition. As the desired resolution is to be improved (increasing s_{max}), however, the computation time will eventually be dominated by that consumed in the second and third steps, which are the same for the PAVA and NACT.

5 DATA SET AND MODEL PARAMETRIZATION

We have collected and processed more than 5000 SH seismograms of 343 earthquakes occurring between 1977 and 1991 recorded at 79 seismographic stations of GDSN/IRIS, GEOSCOPE, CDSN, and RSTN networks with epicentral distance between 15° and 165°. After discarding the data that either are too noisy, or have apparent instrument calibration or timing problems, we use 5041 long-period body-wave records and 1531 long-period mantle-wave records in this study. The data are low-pass filtered with a cut-off frequency 1/32 Hz for body waves and 1/80 Hz for surface waves. We manually select windows in the time domain (see Fig. 3) so that we can use only those portions of data that are associated with major energy
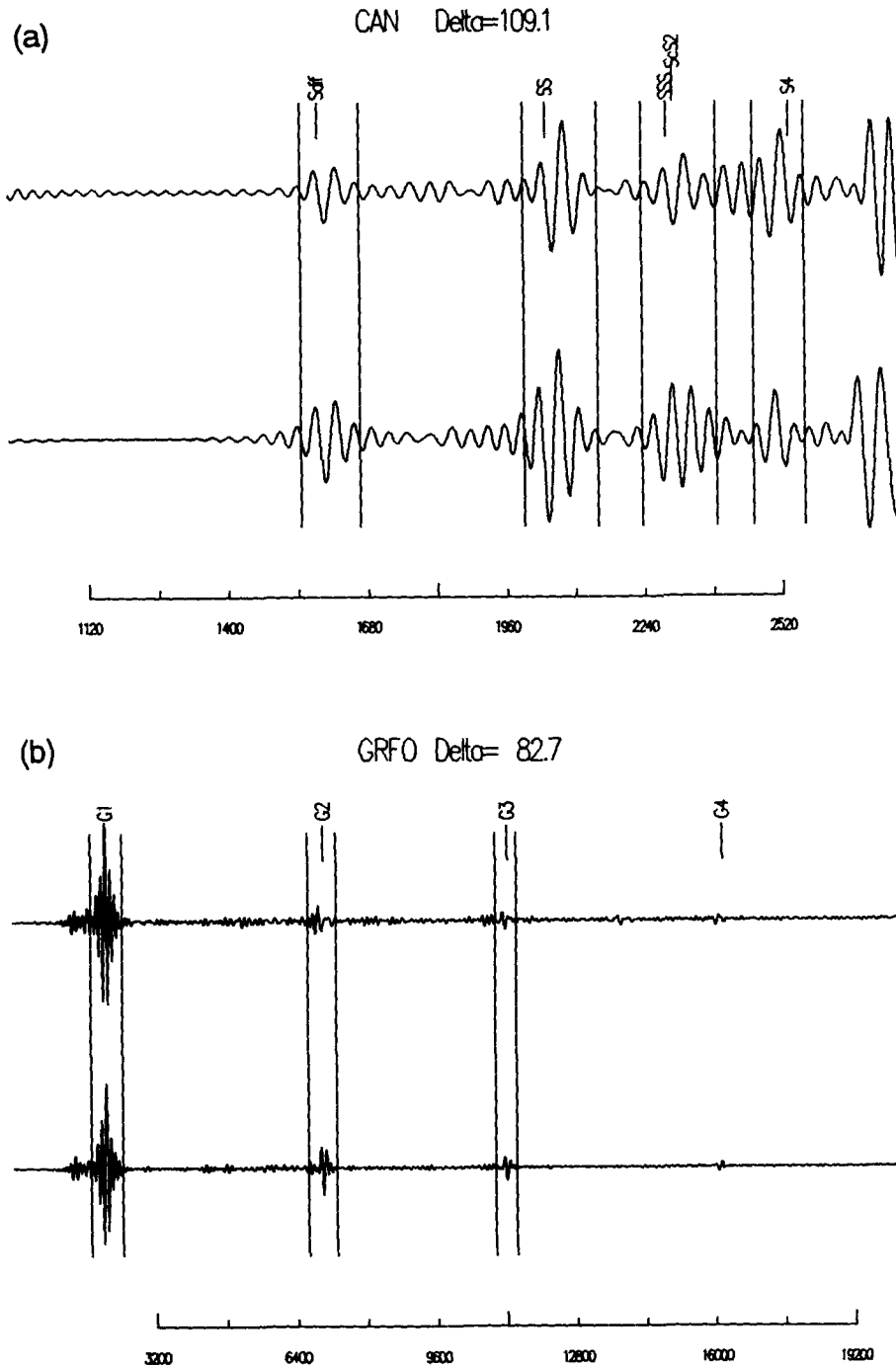
Figure 3. (a) Low-pass filtered ($f < 1/32$ Hz) SH body-wave seismogram of 1988 January 19 Chile earthquake recorded at GEOSCOPE station CAN. (b) Low-pass filtered ($f < 1/80$ Hz) SH mantle-wave seismogram of 1987 April 22 Japan earthquake recorded at GDSN station GRFO. In both cases the observed seismogram is plotted on the top trace, and the synthetic seismogram calculated from the spherical reference model PREM (Dziewonski & Anderson 1981) is plotted on the bottom trace for reference. Only the data within the indicated windows are used in the inversions. The time-scales are in seconds.

arrivals. Obviously this leads to large savings in computation compared with the case in which the whole wave train is used, as in previous studies. Another purpose of windowing is that it gives us flexibility on how to weight different phases in the inversion. For example, if we fit the whole wave train in Fig. 3(a) as a block, the S_{dif} phase will receive a smaller weight in the inversion relative to the SS phase, due to its

smaller amplitude. However, information carried by the S_{dif} phase may be more valuable, considering that the S_{dif} phase samples the bottom of the mantle, which fewer waves reach. The windowing scheme here enables us to assign weights to different phases more appropriately.

We parametrize the model by expanding relative S-velocity perturbations in terms of spherical harmonics Y_{st}

(Edmonds 1960) for the horizontal dependencies and Legendre polynomials f_q for the radial dependency:

$$\frac{\delta v_S(r, \theta, \phi)}{v_S(r)} = \sum_{s=0}^{s_{max}} \sum_{t=-s}^{s} \sum_{q=0}^{q_{max}} c_{stq} Y_{st}(\theta, \phi) f_q(x), \tag{28}$$

where δv_S is the perturbation in S-velocity, $v_S(r)$ is the spherical average S-velocity evaluated at the reference model PREM (Dziewonski & Anderson 1981), r is the radius, θ is the colatitude, ϕ is the longitude, x is the reduced normalized radius (Dziewonski 1984; Woodhouse & Dziewonski 1984), and c_{stq} are coefficients to be determined. For the present experiments, the spherical harmonic expansion is truncated at degree $s_{max} = 8$ and the polynomial expansion is up to degree $q_{max} = 5$ for the upper mantle and up to degree $q_{max} = 7$ for the lower mantle. Thus the total number of unknown coefficients c_{stq} is 1134. Although alternative basis functions may be chosen for the radial dependency (for example Woodward et al. (1993) used Chebyshev polynomials as their basis functions), the choice of basis functions should not have large effects on the comparison of the two inversion approaches studied here.

In the following sections we will show results of resolution analyses and inversions of observed data using the PAVA and NACT, based upon the same weighting scheme and model parametrization as described above. Since weighting schemes and model parametrization are also crucial for inverse problems, we expect that the models developed using the two approaches, respectively, may be similar to each other in the context of the comparison with some other models found in the literature, which are developed using different weighting schemes and model parametrizations in addition to different data and forward theories. On the other hand, the difference between the results using the PAVA and NACT should be purely due to the difference in the theories used.

6 COMPARISON OF FORMAL RESOLUTION RESULTS FROM THE PAVA AND NACT

Since both the PAVA and NACT are non-linear in terms of the relationship between the model and data, the inverse problem has to be solved iteratively. As for formal resolution, however, we choose to perform conventional linear analysis, which may be viewed as investigating to what extent a small perturbation to the final, converged model can be recovered.

For any small model perturbation $\mathbf{m}$, we can express the corresponding data perturbation $\mathbf{d}$ by a linear relationship

$$\mathbf{d} = \mathbf{Am}, \tag{29}$$

where $\mathbf{A}$ is the partial-derivative matrix. Using the synthetic data, we may infer an output model $\mathbf{m}_{out}$ by inversion:

$$\mathbf{m}_{out} = \mathbf{G}(\eta)\mathbf{d}, \tag{30}$$

where $\mathbf{G}(\eta)$ is an inversion matrix with the parameter η representing the characteristics of the a priori data and model variance. For example, $\mathbf{G}(\eta) = (\mathbf{A}^T\mathbf{A} + \eta\mathbf{I})^{-1}\mathbf{A}^T$ is a classical damped least-squares inversion matrix and η in this case is the damping parameter. Substituting eq. (29) into eq.

(30), we obtain

$$\mathbf{m}_{out} = \mathbf{G}(\eta)\mathbf{Am}_{in}, \tag{31}$$

where $\mathbf{G}(\eta)\mathbf{A}$ is called the resolution matrix (Aki & Richards 1980), which depends upon three factors: the dimension and distribution of the data, the theory that governs the calculation of the partial-derivative matrix $\mathbf{A}$, and the a priori information on the data and model variance. Eq. (31) suggests an approach to investigate the resolution of inversion by comparing the output model $\mathbf{m}_{out}$ with a given input model $\mathbf{m}_{in}$. We shall present the results of two parallel experiments using two different theories, the PAVA and NACT, while keeping the data dimension and distribution and parameter η fixed.

In the resolution analyses we simulate the data distribution described in the previous section and use the same damping parameter η as used in inverting the observed data (see next section). We choose the input model to have a checker-board-looking horizontal pattern, obtained using a spherical harmonic degree 8 and order 4 components. The amplitude varies vertically, with sign changes occurring at some depths, as shown in Fig. 4.

Shown in Fig. 5 are vertical cross-section views of the output models using the PAVA (a) and NACT (b). Generally speaking, the resolution decays with depth in both cases due to the fact that fewer phases sample deeper structure. In the lower mantle the result using the NACT is significantly better than that using the PAVA. The amplitude of the former is larger and closer to that of the input model. While the model using the PAVA totally fails to recover the bottom layer (~ 300 km thick) of the input model, the success of NACT is very encouraging. In Fig. 6, maps of the two output models are shown at three depths: 250, 2000, and 2750 km. At a depth of 250 km, the patterns of both models are almost the same as that of the input model, with the amplitude somewhat decayed (by approximately 30 per cent) due to damping. The advantage of NACT becomes obvious at a depth of 2000 km. At a depth of 2750 km, the PAVA is unable to resolve the input structure at all; in contrast, the NACT can still reveal the pattern of the input model very well. Considering that the wavelength of the body waves used here is a few hundreds of kilometres and that the input model varies vertically very rapidly in this region, the deviation between the input and output models is expected for any technique.

Finally, we wish to point out that an excellent result of formal resolution analysis does not prove the reliability of the inversion result, for it is possible to achieve good results in formal resolution analysis with an incorrect forward theory. However, if one theory fails in resolution tests while a more accurate theory performs better, problems exist with the former.

7 DIFFERENCE IN INVERSION RESULTS OBTAINED USING THE PAVA AND NACT

The inversions are performed using the standard damped least-squares method with an a priori starting model in which the hydrostatic ellipticity and crustal thickness are added to the spherically symmetric model PREM (Dziewonski & Anderson 1981). The damping level is determined
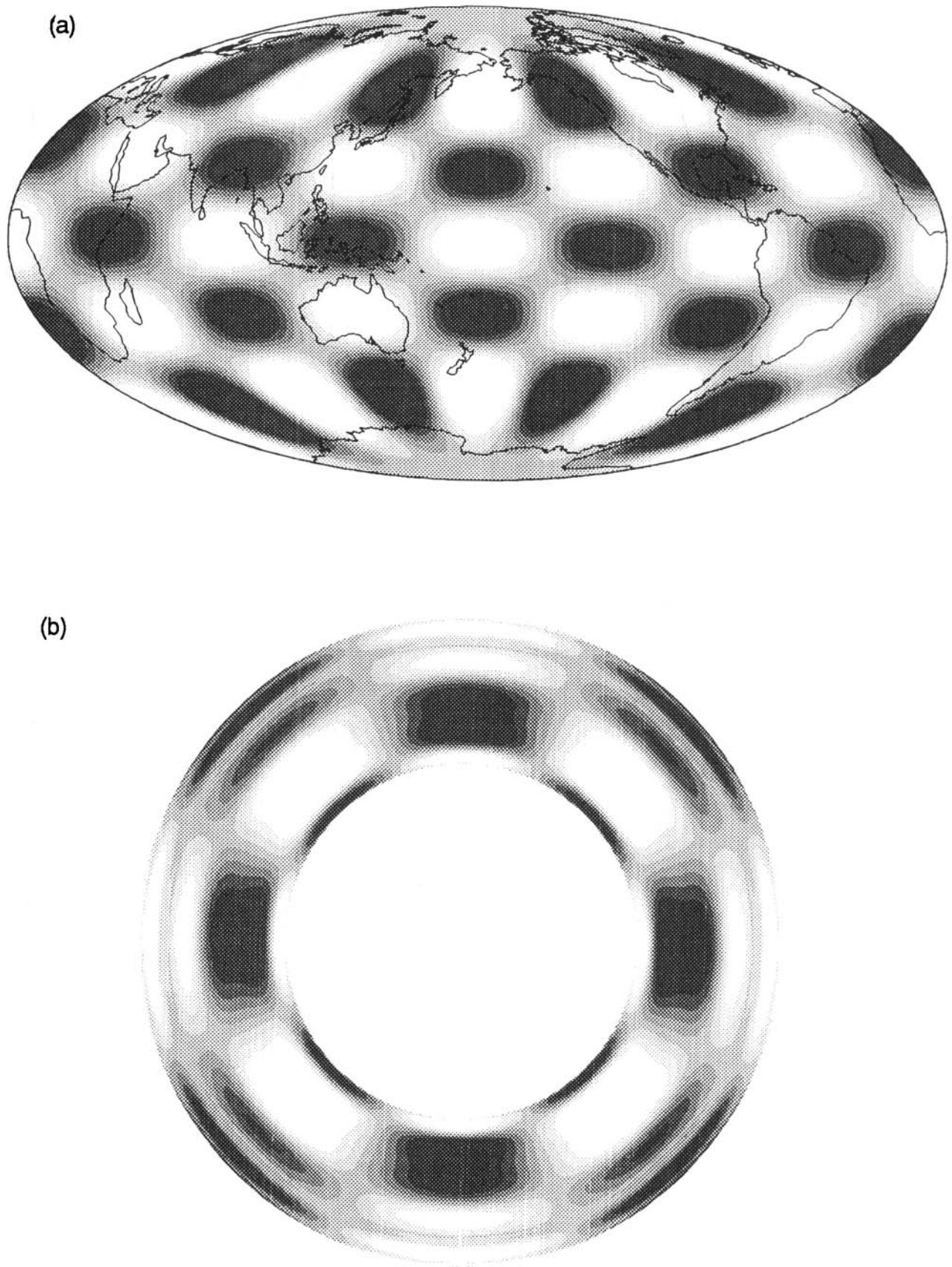
(a)

(b)

Figure 4. (a) A map view of an input model used in resolution analyses. The model has the same horizontal pattern at all the depths, but the amplitude and sign vary vertically. (b) A vertical cross-section view of the input model, cut at the equator of the Earth.
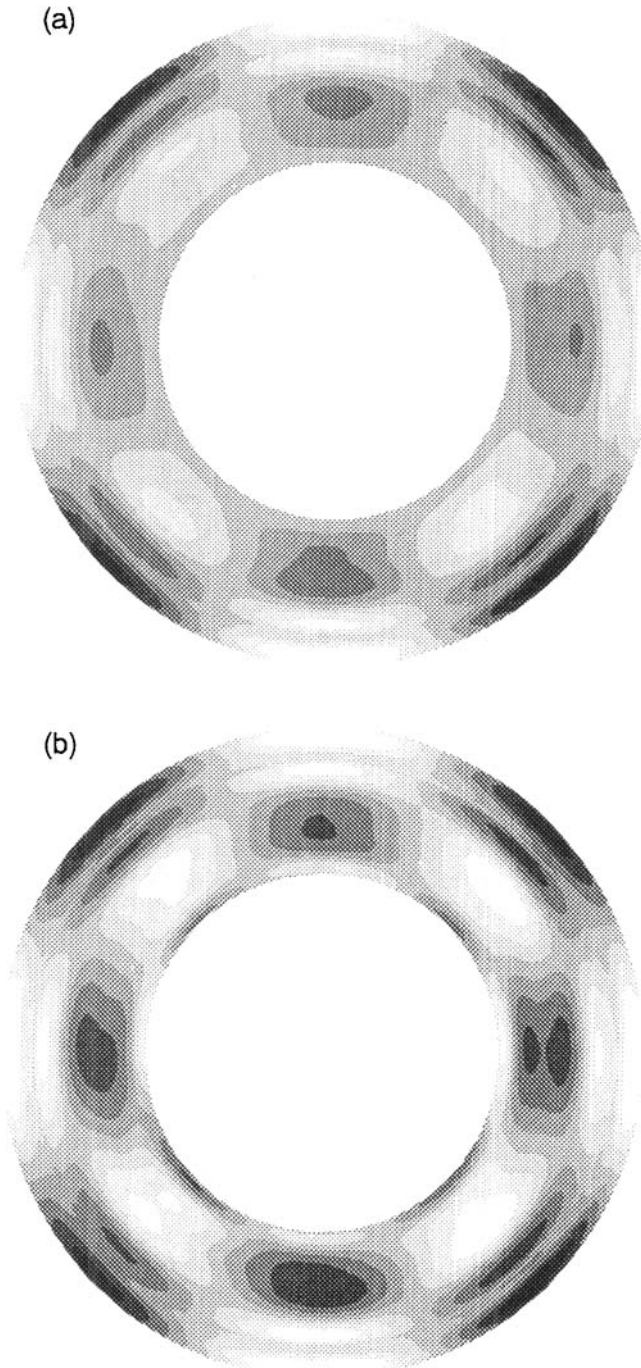
(a)

(b)

Figure 5. (a) A vertical cross-section view of the output model obtained using the PAVA. (b) A vertical cross-section view of the output model obtained using the NACT. For perfect resolution they would look the same as Fig. 4(b).

empirically to balance the trade-off between the data and model variance. The correction for crustal thickness is done in the same manner as in Woodhouse & Dziewonski (1984) with the exception that we replace the ocean–continent distribution by more realistic topographic data (Y. Ricard, personal communication). In the calculation of the partial derivative and synthetic seismograms, the PAVA and NACT are used in parallel and two models (model PAVA and model NACT, respectively) are developed.

A measure of the difference between the two models is given by

$$D \equiv \left(\frac{\sum\limits_{stq} (c_{stq}^{N} - c_{stq}^{P})^2}{\sqrt{\sum\limits_{stq} (c_{stq}^{N})^2 \sum\limits_{stq} (c_{stq}^{P})^2}} \right)^{1/2}, \tag{32}$$

where c_{stq}^{N} and c_{stq}^{P} are the c_{stq} coefficients (see eq. 28) of models NACT and PAVA, respectively. The D value is 31 per cent for the upper mantle and 55 per cent for the lower mantle. The difference between the two models is large. We believe that the reason for the better agreement in the upper mantle is twofold:

(1) the two theories differ less in the upper mantle, as sensitivity kernels suggest;

(2) the data coverage is better in the upper mantle, which could reduce the bias caused by horizontal averaging of the PAVA method.

This increase in difference between models PAVA and NACT as a function of depth is also shown in Fig. 7, in which the rms lateral heterogeneity as a function of depth is plotted for models PAVA and NACT and the differential model DIFF (model NACT minus model PAVA). At the shallow depths, the rms of model DIFF is smaller than those of PAVA and NACT, while in the deep mantle it is as large as those of PAVA and NACT.

It is also of interest to look at the spherical harmonic spectra of the differential model DIFF. The spherical harmonic (rms amplitude) spectrum C as a function of angular degree s is defined by

$$C(s) = \left(\sum_{t=-s}^{s} \sum_{q=0}^{q_{max}} c_{stq}^{2} \right)^{1/2}. \tag{33}$$

The spectra of model DIFF are shown in Fig. 8 for the upper and lower mantles separately, along with the spectra of models PAVA and NACT. The amplitude of model DIFF increases with spherical harmonic degree. This is not surprising, since the bias due to the averaging operation by the PAVA should be more severe for higher degree structure.

Finally, it is interesting to see how the two models are correlated with some independent geophysical observations. One of the most important surface observations, which reflects the Earth's interior physical properties, is non-hydrostatic geoid. According to the geodynamic theory of Richards & Hager (1984), the geoid $g(\theta, \phi)$ may be expressed, in a first-order approximation, as a function of the 3-D distribution of density anomaly $\delta\rho(r,\theta, \phi)$ and spherical average viscosity $v(r)$:

$$g(\theta, \phi) = f(\delta\rho(r, \theta, \phi), v(r)). \tag{34}$$

As an approximation we assume that the density anomaly can be obtained from

$$\frac{\delta\rho(r, \theta, \phi)}{\rho(r)} = c(r) \frac{\delta v_S(r, \theta, \phi)}{v_S(r)}, \tag{35}$$

where $c(r)$ is a scaling factor, $\rho(r)$ and $v_S(r)$ are the density and S-velocity of a spherical reference model (PREM of Dziewonski & Anderson 1981), respectively, and
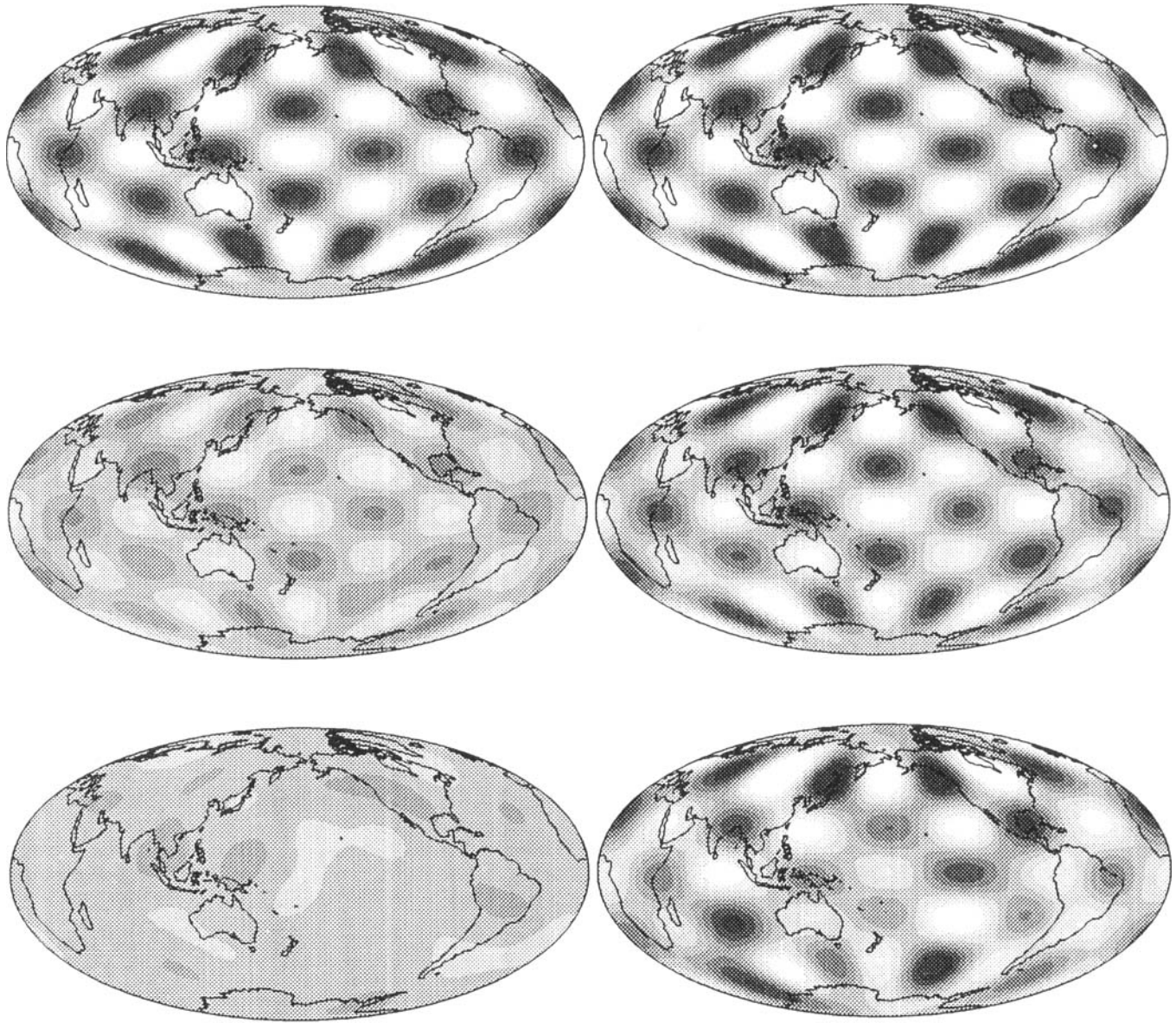
Figure 6. Maps of output models at depths of 250 km (top), 2000 km (middle) and 2750 km (bottom). The maps shown in the left-hand column are obtained using the PAVA and those shown in the right-hand column are the results from the NACT. The scales are normalized for each depth so that for perfect resolution, a map would look the same as the one in Fig. 4(a), with the exception that the contour interval of the map for the PAVA at 2750 km (left bottom) is half of that described here.

$\delta v_S(r, \theta, \phi)$ is inferred from seismic data. Our knowledge on $v(r)$ and $c(r)$ is poor. Among others, Corrieu, Ricard & Froidevaux (1994) have developed a procedure to infer the viscosity distribution $v(r)$ and the scaling factor $c(r)$ by optimizing the fit of the synthetic geoid to the observations. Since the observed geoid has a structure dominated by spherical harmonic degrees 2 and 3 (87 per cent of the total power between degrees 2 and 8 is from degrees 2 and 3), the distributions $v(r)$ and $c(r)$ inferred in this manner are constrained very largely by the degree 2 and 3 structure of the seismic model. The synthetic geoid of higher degrees (degrees 4–8) obtained using these $v(r)$ and $c(r)$ can, therefore, be used as means of assessing the seismic models. Using a program modified from the one by Corrieu *et al.* (1994) and assuming whole-mantle convection, we obtain

the optimized $v(r)$ and $c(r)$ for models PAVA and NACT. The results are shown in Fig. 9. The two models yield similar $v(r)$ and $c(r)$ and can predict good variance reductions in fitting the total geoid field (76.0 per cent from PAVA and 78.5 per cent from NACT). This reflects the fact that the two models are very similar in their very low-degree structure. However, if we look at higher degrees, the difference between the two models shows up. In Fig. 10, the correlation coefficients between the synthetic and observed geoid are demonstrated for each degree. The synthetic geoid predicted by model NACT correlates with the observed better in degrees 4, 5, and 7 than that predicted by model PAVA. For degree 6, both models predict excellent correlation coefficients. The failures for degree 8 of both models might be associated with the truncation of spherical
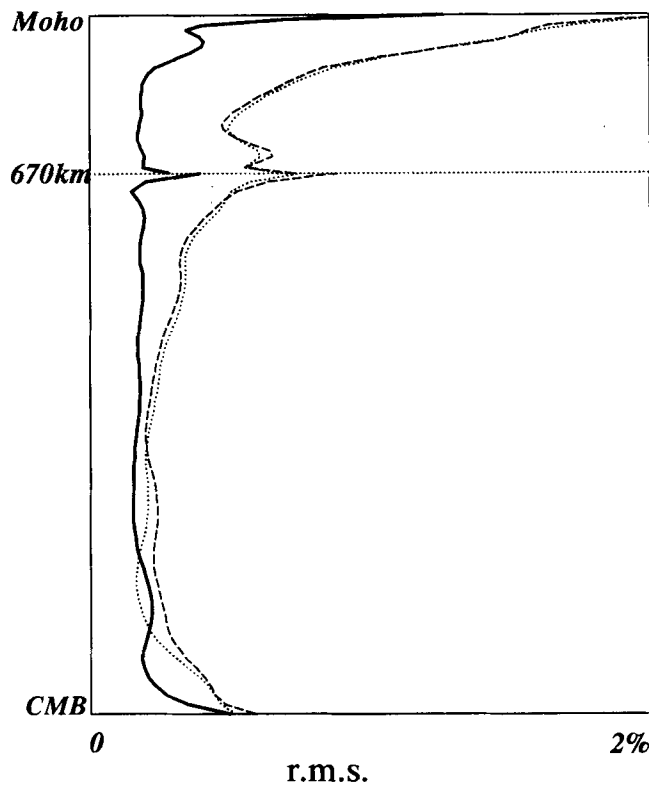
Figure 7. The vertical distributions of rms lateral heterogeneity of model NACT (dashed curve), model PAVA (dotted curve), and the differential model (model NACT minus model PAVA, solid curve). The spiky peaks, with widths of approximately 80 km, at the 670 km discontinuity can be attributed to the so-called edge effect of Legendre polynomials (e.g. Dziewonski *et al.* 1993).

harmonic expansion at that degree. Nevertheless, the experiment shown here indicates clearly that model NACT predicts the geoid significantly better than model PAVA.

8   CONCLUSIONS

The traditional waveform inversion technique, PAVA, has been used widely in global seismic tomography. In this study we have presented a new inversion technique, NACT, based upon a more accurate asymptotic theory, which is able to bring out the ray character of body waves by taking cross-branch modal coupling into account. We have demonstrated that the new technique has higher resolution power. From a given data set of reasonable size, two models are inferred using the traditional PAVA and the new NACT, respectively. The results are significantly different and the one obtained using the NACT can predict the surface geoid, an independent observation, better than that obtained using the PAVA. The increase in computation time associated with replacing the PAVA by the NACT is manageable. In addition, the relative difference in computation time decreases as more model parameters are introduced.

Based upon the advantage of the NACT as shown in this study, we are very optimistic that the new technique has
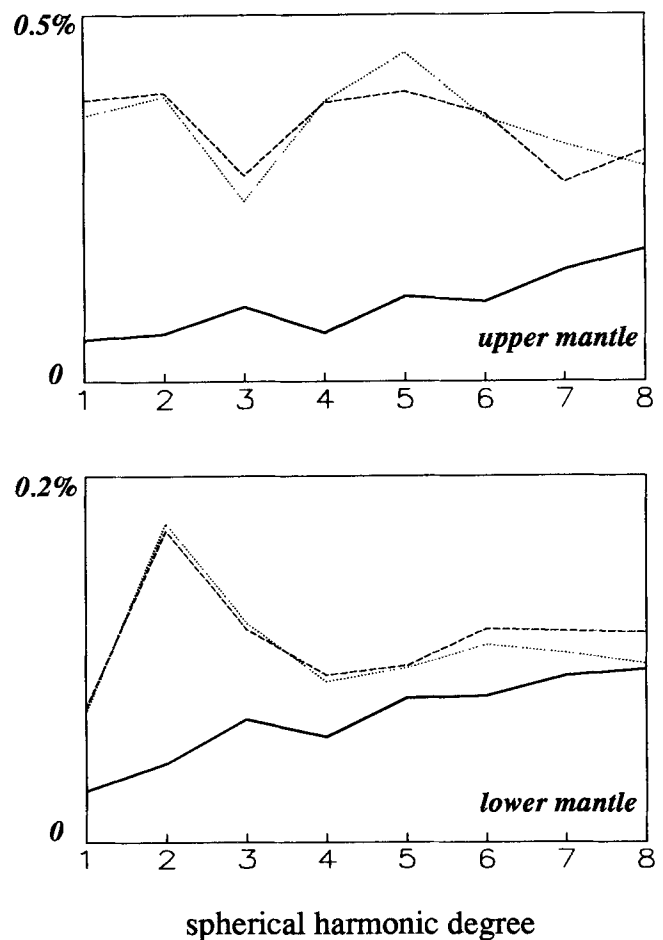

spherical harmonic degree

Figure 8. The amplitude spectra of model NACT (dashed curve), model PAVA (dotted curve), and the differential model (model NACT minus model PAVA, solid curve) in the upper and lower mantles.

great power for developing more reliable, higher resolution models. We are working on applying this technique to a data set larger than the one used here and a model of higher resolution will be available shortly, for which the level of horizontal detail will make it directly comparable with recently published whole-mantle models (Masters *et al.* 1992; Su *et al.* 1994). The next challenge will be to incorporate off-great-circle effects, which can be modelled using higher order asymptotics (Park 1987; Romanowicz 1987). Modelling amplitude anomalies caused by elastic focusing of waves due to off-great-circle structure is very important in retrieving anelastic mantle structure (e.g. Romanowicz 1990, 1994; Durek, Ritzwoller & Woodhouse 1993).

ACKNOWLEDGMENTS

We are very grateful to Corrieu, Ricard & Froidevaux for their program for calculating the synthetic geoid. Yves Le Stunff has kindly spent time teaching us how to use this program. The revision of the manuscript has benefited from very helpful comments by Robert Geller. We also thank Adam Dziewonski and Wei-jia Su who provided us with results from the Harvard seismology group before
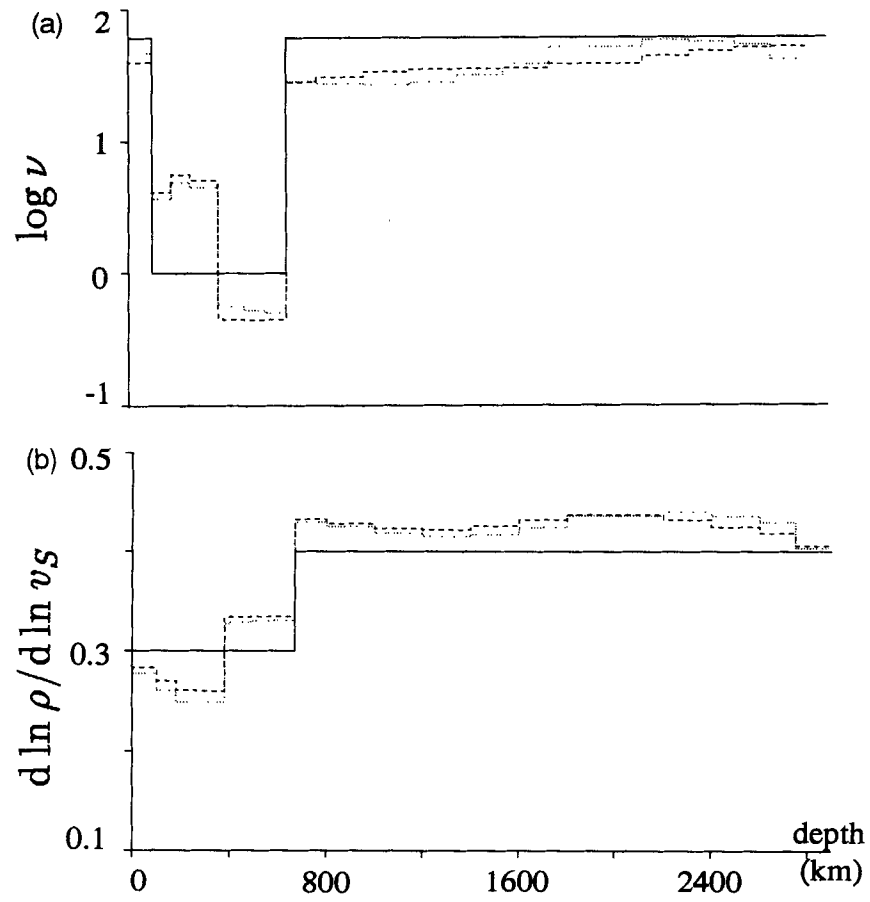
Figure 9. (a) The optimized viscosity $v(r)$ and (b) scaling factor $c(r)$. In both diagrams the solid line represents the *a priori* starting model towards which damping is applied; the dashed line represents the result from model NACT; and the dotted line represents the result from model PAVA.
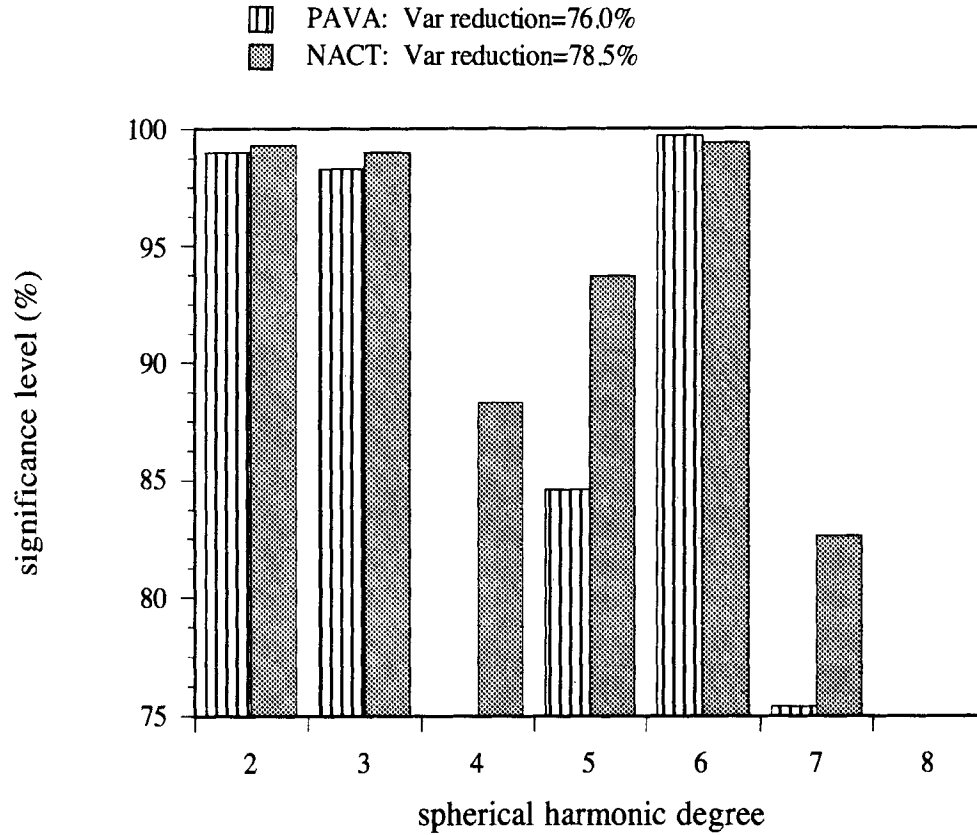
⊞ PAVA: Var reduction=76.0%
▨ NACT: Var reduction=78.5%

Figure 10. Correlation coefficients presented as significance level between the observed geoid and synthetic geoid.

REFERENCES

Aki, K. & Richards, P.G., 1980. Quantitative Seismology: Theory and Methods, W.H. Freeman, San Francisco.

Corrieu, V., Ricard, Y. & Froidevaux, C., 1994. Converting mantle tomography into mass anomalies to predict the Earth's radial viscosity, Phys. Earth planet. Inter., 84, 3-13.

Dahlen, F.A., 1987. Multiplet coupling and the calculation of synthetic long-period seismograms, Geophys. J. R. astr. Soc., 91, 241-254.

Durek, J.J., Ritzwoller, M.H. & Woodhouse, J.H., 1993. Constraining upper mantle anelasticity using surface wave amplitude anomalies, Geophys. J. Int., 114, 249-272.

Dziewonski, A.M., 1984. Mapping the lower mantle: Determination of lateral heterogeneity in P velocity up to degree and order 6, J. geophys. Res., 89, 5929-5952.

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

Dziewonski, A.M. & Woodward, R.L., 1992. Acoustic imaging at the planetary scale, in Acoustical Imaging, Vol. 19, pp. 785-797, eds Ermert, H. & Harjes, H.-P., Plenum, New York.

Dziewonski, A.M., Forte, A.M., Su, W.-J. & Woodward, R.L., 1993. Seismic tomography and geodynamics, in Relating Geophysical Structures and Process: Proceedings of the Jeffreys Symposium, Geophys. Monogr. Ser., Vol. 76, pp. 67-105, eds Aki, K. & Dmowska, R., Am. geophys. Un., Washington, D.C.

Edmonds, A.R., 1960. Angular Momentum and Quantum Mechanics, Princeton University Press, Princeton, NJ.

Geller, R.J. & Hara, T., 1993. Two efficient algorithms for iterative linearized inversion of seismic waveform data, Geophys. J. Int., 115, 699-710.

Gilbert, F. & Dziewonski, A.M., 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, Phils. Trans. R. Soc. Lond., A, 278, 187-269.

Hara, T., Tsuboi, S. & Geller, R.J., 1991. Inversion for laterally heterogeneous earth structure using a laterally heterogeneous starting model: preliminary results, Geophys. J. Int., 104, 523-540.

Hara, T., Tsuboi, S. & Geller, R.J., 1993. Inversion for laterally heterogeneous upper mantle S-wave velocity structure using iterative waveform inversion, Geophys. J. Int., 115, 667-698.

Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y., 1990. Whole mantle P-wave travel time tomography, Phys. Earth planet. Inter., 59, 294-328.

Jordan, T.H., 1978. A procedure for estimating lateral variations from low-frequency eigenspectra data, Geophys. J. R. astr. Soc., 52, 441-455.

Li, X.-D., 1990. The asphericity of the Earth from free oscillations, PhD thesis, Harvard University, Cambridge, MA.

Li, X.-D. & Tanimoto, T., 1993a. Waveforms of long-period body waves in a slightly aspherical Earth model, Geophys. J. Int., 112, 92-102.

Li, X.-D. & Tanimoto, T., 1993b. Waveform inversion of long-period seismic data for global structure, in Seismic Tomography: Theory and Practice, pp. 64-91, eds Iyer, H.M. & Hirahara, K., Chapman & Hall, London.

Liu, X-f., Su, W.-J. & Dziewonski, A.M., 1994. Improved resolution of the lowermost mantle shear wave velocity structure obtained using SKS-S data, EOS Trans. Am. geophys. Un., 75, Spring Meeting Suppl., 232.

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

Masters, G., Bolton, H. & Shearer, P., 1992. Large-scale 3-dimensional structure of the mantle, EOS Trans. Am. geophys. Un. 73, Spring Meeting Suppl., 201.

Mochizuki, E., 1986a. Free oscillations and surface waves of an aspherical Earth, Geophys. Res. Lett., 13, 1478-1481.

Mochizuki, E., 1986b. The free oscillations of an anisotropic and heterogeneous Earth, Geophys. J. R. astr. Soc., 86, 167-176.

Park, J., 1987. Asymptotic coupled-mode expressions for multiplet amplitude anomalies and frequency shifts on an aspherical Earth, Geophys. J. R. astr. Soc., 90, 129-169.

Park, J., 1990. The subspace projection method for constructing coupled-mode synthetic seismograms, Geophys. J. Int., 101, 111-123.

Pulliam, R.J., Vasco, D.W. & Johnson, L.R., 1993. Tomographic inversions for mantle P-wave velocity structure based on the minimization of l_2 and l_1 norms of ISC travel time residuals, J. geophys. Res., 98, 699-734.

Richards, M.A. & Hager, B.H., 1984. Geoid anomalies in a dynamic Earth, J. geophys. Res., 89, 5987-6002.

Romanowicz, B., 1987. Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the Earth's normal modes, Geophys. J. R. astr. Soc., 90, 75-100.

Romanowicz, B., 1990. The upper mantle degree 2: Constraints and inferences from global mantle wave attenuation measurements, J. geophys. Res., 95, 11 051-11 071.

Romanowicz, B., 1994. Anelastic tomography: a new perspective on upper-mantle thermal structure, Earth Planet. Sci. Lett., 128, 113-121.

Romanowicz, B. & Snieder, R., 1988. A new formalism for the effect of lateral heterogeneity on normal modes and surface waves—II. General anisotropic perturbation, Geophys. J. Int., 93, 91-89.

Stack, P.B. & Nikolayev, D.I., 1993. Toward tubular tomography, J. geophys. Res., 98, 8095-8106.

Su, W.-J. & Dziewonski, A.M., 1991. Predominance of long-wavelength heterogeneity in the mantle, Nature, 352, 121-126.

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., 1986a. The Backus–Gilbert approach to the three-dimensional structure in the upper mantle, II. SH and SV velocity, Geophys. J. R. astr. Soc., 84, 49-70.

Tanimoto, T., 1986b. Free oscillations in a slightly anisotropic Earth, Geophys. J.R. astr. Soc., 87, 493-517.

Tanimoto, T., 1990. Long-wavelength S-wave velocity structure throughout the mantle, Geophys. J. Int., 100, 327-336.

Um, J., Dahlen, F.A. & Park, J., 1991. Normal mode multiplet coupling along dispersion branch, Geophys. J. Int., 106, 11-36.

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., 1983. The joint inversion of seismic waveforms for lateral variations in Earth structure and earthquake source parameters, eds Kanamori, H. & Boschi, E., Proc. 'Enrico Fermi' Int. Sch. Phys., LXXXV, 366-397.

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. & Dziewonski, A.M., 1984. Mapping the upper mantle: Three dimensional modelling of Earth structure by inversion of seismic waveforms, J. geophys. Res., 89, 5953-5986.

Woodward, R.L. & Masters, G., 1991a. Global upper mantle structure from long-period differential travel-times, J. geophys. Res., 96, 6351-6377.

Woodward, R.L. & Masters, G., 1991b. Lower-mantle structure from ScS-S differential travel-times, Nature, 352, 231-233.

Woodward, R.L., Forte, A.M., Su, W.-J. & Dziewonski, A.M.,

1993. Constraints on the large-scale structure of the Earth's mantle, in *Evolution of the Earth and Planets, Geophys. Monogr. Ser.*, Vol. 74, pp. 89–109, eds Takahashi, E., Jeanloz, R. & Rubie, D., AGU, Washington, DC.

Zhang, Y.-S. & Tanimoto, T., 1993. High-resolution global upper mantle structure and plate tectonics, *J. geophys. Res.*, **98**, 9793–9823.

APPENDIX A

Defining

$$g_l^{(1)}(\Delta) = \sqrt{\frac{2}{\lambda \pi \sin \Delta}} \cos\left(\lambda \Delta - \frac{\pi}{4}\right), \tag{A1}$$

$$g_l^{(2)}(\Delta) = \sqrt{\frac{2}{\lambda \pi \sin \Delta}} \sin\left(\lambda \Delta - \frac{\pi}{4}\right), \tag{A2}$$

$$T_{kk'}^{(1)} = \sum_{N=-1}^{1} \sum_{M=-2}^{2} i^{N+M} R_{kN} S_{k'M} \cos\left(\frac{N+M}{2}\pi\right), \tag{A3}$$

$$T_{kk'}^{(2)} = \sum_{N=-1}^{1} \sum_{M=-2}^{2} i^{N+M} R_{kN} S_{k'M} \sin\left(\frac{N+M}{2}\pi\right), \tag{A4}$$

where $\lambda \equiv l + 1/2$ and R_{kN} and $S_{k'M}$ are defined in eqs (26) and (27) in Li & Tanimoto (1993a), we may express $Q_{kk'}^{(1)}$ and $Q_{kk'}^{(2)}$ in eq. (13) as

$$Q_{kk'}^{(1)} = (1 - I_{kk'}/2)[P_{kk'}^{(1)} + P_{k'k}^{(1)}], \tag{A5}$$

and

$$Q_{kk'}^{(2)} = -(1 - I_{kk'}/2)[P_{kk'}^{(2)} - P_{k'k}^{(2)}], \tag{A6}$$

where

$$P_{kk'}^{(1)} = g_l^{(1)}(\Delta) T_{kk'}^{(1)} - g_l^{(2)}(\Delta) T_{kk'}^{(2)}, \tag{A7}$$

and

$$P_{kk'}^{(2)} = g_l^{(1)}(\Delta) T_{kk'}^{(2)} + g_l^{(2)}(\Delta) T_{kk'}^{(1)}. \tag{A8}$$

APPENDIX B

Both coefficients $A_{kk'}^q$ and $B_{kk'}^q$ in eq. (18) may be factorized into three factors:

$$A_{kk'}^q = M_{kk'}^q \cdot U_{kk'} \cdot D_{kk'}(\tau; \hat{\omega}_k, \hat{\omega}_{k'}), \tag{B1}$$

and

$$B_{kk'}^q = M_{kk'}^q \cdot V_{kk'} \cdot D_{kk'}(\tau; \hat{\omega}_k, \hat{\omega}_{k'}), \tag{B2}$$

where

$$M_{kk'}^q = \int_0^a M_{kk'}(r) \cdot \mathbf{f}_q r^2 \, dr \tag{B3}$$

depend only upon the reference spherical earth model and the choice of basis functions $\mathbf{f}_q$. $U_{kk'}$ and $V_{kk'}$ are time-independent coefficients describing the source and receiver effects and can be given in terms of $Q_{kk'}^{(1)}$ and $Q_{kk'}^{(2)}$ (see eqs A5 and A6) as

$$U_{kk'} = \begin{cases} \frac{1}{2} Q_{kk'}^{(1)} & \text{if } j \neq 0 \\ Q_{kk'}^{(1)} & \text{if } j = 0 \end{cases}, \tag{B4}$$

$$V_{kk'} = \begin{cases} \frac{1}{2} Q_{kk'}^{(2)} & \text{if } j > 0 \\ -\frac{1}{2} Q_{kk'}^{(2)} & \text{if } j < 0, \\ 0 & \text{if } j = 0 \end{cases} \tag{B5}$$

where $j = l' - l$ is the difference between the angular degrees of multiplets k' and k. $D_{kk'}$ are given in eq. (12).